ABSTRACT BOOK

JSM 2008

Denver, Colorado
August 3–7, 2008

ABSTRACT BOOK

JSM
2008

Denver, Colorado
August 3–7, 2008

Table of Contents

Session Abstracts

▲ Themed Session ● Applied Session ✱ Presenter

1 Designing Courses on Statistical Computing

Section on Statistical Computing, Section on Statistical Education, WNAR
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Computing in the Nonintroductory Statistics Classes

✱Deborah Nolan, University of California, Berkeley, Dept of Statistics, 367 Evans Hall MC 3860, Berkeley, CA 94720-3860, *nolan@stat.berkeley.edu*

Key Words: statistical computing, undergraduate program, curriculum development

Our current undergraduate curriculum often omits the important area of computational literacy and data technologies which is becoming increasingly more important for all aspects of the practice of data analysis and statistics. We will outline approaches to bringing these topics into the statistics curriculum, either by introducing new focused classes or by integrating topics into existing classes. We will discuss our experiences in doing this over the previous four years. We firmly believe that essential computational skills provides an enormous value-added to our students and excites and empowers them to be able to approach data problems in an involved, fun and creative rather than in a formulaic, old-fashioned manner.

Computational Graphics

✱Leland Wilkinson, Northwestern University, Chicago, IL , *leland.wilkinson@gmail.com*

Key Words: graphics, graduate courses, statistical computing

Statistical graphics are at the core of statistical practice, yet few graduate programs have integrated graphics into their graduate curriculum. There are two approaches to introducing graduate students to the field of statistical graphics—one involving a semester-long course, and the other involving a two-week segment of a graduate course in statistical computing or data analysis. The former approach should include a broader view of information visualization, as well as an introduction to design and use of graphical software. The latter approach necessarily involves a minimal introduction to graphics practice and available software. This talk will discuss potential curricula for both approaches.

Computing in the Graduate Curriculum

✱Duncan Temple Lang, University of California, Davis, Department of Statistics, 4210 Math Sciences Bldg., 1 Shields Avenue, Davis, CA 95616, *duncan@wald.ucdavis.edu*

Key Words: Computing, curriculum, graduate, technologies

Our graduate students need to be able to work with data from a myriad of sources, in different formats and in large volume. They need to work with code developed by other people. And they routinely need to perform complex and intensive simulation studies. Modern computing, along with more traditional numerically-oriented techniques (e.g., optimization), is fast becoming a key skill in academic and industrial work, yet our curricula have hardly changed during the computer "revolution." Our students need to be mathematically and computationally literate. I will discuss possible topics and syllabi for integrating computing into the Master's and PhD programs, focusing on the different types of students we see in statistics today and prioritizing these topics and relative to other competing statistical topics in our curricula. This dimension should strengthen our discipline.

2 Recent Advances in Computer Experiments: New Methods with Diverse Applications ●

Section on Physical and Engineering Sciences, Section on Quality and Productivity
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Gaussian Process Models for Computer Experiments with Qualitative and Quantitative Factors

✱C.F. Jeff Wu, Georgia Institute of Technology, School of Industrial and Systems Engineering, Atlanta, GA 30332, *jeffwu@isye.gatech.edu*; Zhiguang (Peter) Qian, University of Wisconsin-Madison; Huaiqing Wu, Iowa State University

Key Words: Gaussian process, computer experiments, design of experiments, Information technology, optimization, spatial process

Modeling experiments with qualitative and quantitative factors is an important issue in computer modeling. In this work a framework for building Gaussian process models that incorporate both qualitative and quantitative factors is proposed. The key to the development of these new models is an approach for constructing correlation functions with qualitative and quantitative factors. An iterative estimation procedure is developed for the proposed models. Modern optimization techniques are used in the estimation to ensure the validity of the constructed correlation functions. The proposed method is illustrated with a real example from the IT industry for modeling enterprise data center thermal distribution.

A Bayesian Approach for the Identification of Pollution Source Directions

✱C. Shane Reese, Brigham Young University, Department of Statistics, 230 TMCB, Provo, UT 84602, *reese@stat.byu.edu*

Key Words: computer experiments, Bayesian methods, Markov chain Monte Carlo, Reversible Jump MCMC, Pollution source apportionment, Environmental statistics

Pollution source apportionment (PSA) is the practice of identifying and describing pollution sources and their contributions. PSA frequently requires the identification of source directions, often as a post-analysis check to

ensure that the contribution estimates are reasonable. This paper develops a method of identifying source directions using Bayesian regression. MCMC is used to evaluate the complex relationship among observed pollutant concentrations, available meteorological information, and unknown source direction parameters. The method is flexible enough to identify multiple source directions for cases in which a species or source type of interest is emitted at more than one location, and Reversible Jump MCMC is used to evaluate the appropriate number of sources.

Calibration and Ensemble Prediction for Multiple Computer Models and Data Sources

✱Brian Williams, Los Alamos National Laboratory, Statistical Sciences Group, Los Alamos National Lab, Los Alamos National Lab, NM 87545, *brian@lanl.gov*; Thomas J. Santner, The Ohio State University; Max Morris, Iowa State University

Key Words: Computer experiments, Gaussian process, Multivariate, Bayesian methods, Model averaging

With the proliferation of mathematical models and numerical methods for their solution, we are often faced with several computer models that can be used to predict the same outcome. While these models have common engineering inputs, their different physics or biological simplifications often lead to model-specific calibration parameters. In addition, some applications allow multiple data sources which allow calibration of subsets of the common parameters in one or more of the computational models. This talk will describe methodology that uses simultaneously the outputs from all available computer models and data sources to calibrate unknown model parameters and to provide ensemble predictions of the code output. An example will be given to illustrate the proposed methods.

Design and Analysis of Computer Experiments in High Dimensions

✱Derek Bingham, Simon Fraser University, Department of Statistics and Actuarial Science, 8888 University Drive, Burnaby, International V5A 1S6 Canada, *dbingham@stat.sfu.ca*

Key Words: Gaussian process, computer experiments, design of experiments, likelihood methods, high-dimensional data, cosmology

Gaussian process (GP) models have proven to be very useful in modeling computer simulation output. This is because many computer codes are essentially noiseless and respond very smoothly to changes in input settings. In some applications, the simulation output is a function of a large p-dimensional input vector x. After a fixed number of simulations are carried out, a GP model can be used to predict the simulation output at untried settings. It turns out, that the run sizes required for high dimensional inputs can be computationally burdensome for fitting the GP. In this talk, we describe strategies we've found useful for fitting such models when the number of simulation runs is large. This work is motivated by an ongoing application in cosmology.

3 Statistical Issues for Biodiversity Assessments ●

Section on Statistics and the Environment
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Predicting Seasonal Changes in Species Composition and Richness While Accounting for Imperfect Detectability

✱Robert M. Dorazio, U.S. Geological Survey/University of Florida, Department of Statistics, University of Florida, Gainesville, FL 32611-0339, *bdorazio@ufl.edu*; Marc Kery, Swiss Ornithological Institute; J. Andrew Royle, U.S. Geological Survey

Key Words: biodiversity, species richness, open community, dynamic model

Species richness is perhaps the most widely used metric of biodiversity. However, in surveys of biological communities, some species typically go undetected. Therefore, statistical models are needed to estimate community-level characteristics, such as richness, while accounting for the imperfect detectability of individual species. Imperfect detection is also important when trying to infer the seasonal dynamics of communities, which often include the superposition of species-specific phenologies. For example, seasonal periods of activity (for animals) or growth (for plants) often vary among species. We develop a framework for inferring seasonal changes in species composition and richness based on elemental models of species- and site-specific occurrence. This framework is applied in an analysis of the seasonal dynamics of butterfly species in Switzerland.

Modeling Mixtures of Three States of a Count Process: A Zero-State and Two Poisson Count States

✱Mary C. Christman, University of Florida, Department of Statistics, IFAS PO Box 110339, Gainsville, FL 32611, *mcxman@ufl.edu*

Key Words: Poisson Process, spatial models, zero-inflation

In conservation biology it is important to identify spatial regions with both high biodiversity and endemics (species found only in one place). When species richness is low, usually due to detectability issues, the unobserved true diversity needs to be predicted in order to develop appropriate management plans. We develop an extension to the zero-inflated Poisson regression model that simultaneously models the counts as a mixture of three processes: a zero state, a Poisson count state for the majority of observations, and a second Poisson count state for the very high species richness counts. The estimators of the regression coefficients are obtained using a Bayesian hierarchical approach with non-informative priors. We show that this approach has better predictive capability than the two-part zero-inflated count model and elucidates the underlying mechanisms determining diversity.

A Hierarchical Spatial Model for Predicting Forest Biodiversity Across Large Domains

✱Andrew O. Finley, Michigan State University, Department of Forestry and Department of Geography, Natural Resources Building, East Lansing, MI 48824-1222, *finleya@msu.edu*; Sudipto Banerjee, The University of Minnesota; Ronald E. McRoberts, U.S. Forest Service

Key Words: hierarchical, Bayesian, predictive process, MCMC, biodiversity, spatial

Predictions of forest biodiversity indices at a fine resolution across large regions are invaluable to forest management and conservation initiatives.

We model several important indices of forest biodiversity (e.g., species and stand structure richness) using a data set consisting of >50,000 observations across the north-central US. The proposed analysis employs univariate and multivariate predictive process models with spatially adaptive slopes to address issues of nonstationarity and anisotropy. These models adopt a process-based reduced-rank kriging approach that operates on a lower-dimensional subspace thereby reducing the computational burden of these complex models. In addition to the use of the predictive process, we use tapering positive definite functions to define model covariance matrices, which permit the use of sparse linear systems for efficient matrix computations.

Predicting and Mapping Forest Species Compostion and Structural Attributes

✱Ronald E. McRoberts, U.S. Forest Service, 1992 Folwell Avenue, St. Paul, MN 55108, *rmcroberts@fs.fed.us*

Key Words: nearest neighbors, multinomial logistic regression, discriminant analysis, satellite imagery

Nonparametric, multivariate nearest neighbors techniques have been shown to be useful for simultaneously predicting multiple forest attributes from forest inventory and satellite image data. A two-step nearest neighbors algorithm is proposed in which the class of a relevant categorical variable such as land cover is predicted in the first step, and continuous variables such as tree volume are predicted in the second step subject to the constraint that all nearest neighbors must come from the predicted class of the categorical variable. Nearest neighbors, multinomial logistic regression, and discriminant analysis techniques are compared for use in the first step, whereas only the k-Nearest Neighbors (k-NN) technique are used in the second step because most forest inventory variables do not satisfy the distributional assumptions necessary for parametric multivariate techniques.

4 Statistics in Defense and National Security: Past, Present, and Future ▲

Section on Statisticians in Defense and National Security, Social Statistics Section, *CHANCE*
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Statistics and National Defense Until 1960

✱Fritz Scheuren, The University of Chicago, NORC, 1402 Ruffner Rd, Alexandria, VA 22302, *scheuren@aol.com*

The American Statistical Association, especially at its inception, emphasized the first word in the organization's name, being founded in 1839, among other reasons, to support the 1840 Census. No such lineage and linkage between ASA and national defense has been traced so far to a specific past 19th Century event. Nonetheless, while ASA was set up primarily for peaceful pursuits, it played a wartime role as science and statistical thinking increasingly permeated government decisionmaking in the 19th and early 20th Century. The emphasis in the paper will be on such figures as Florence Nightingale, Solomon Kullback, Mort Kupperman, Sam Wilks, and Abraham Wald. I will leave to other speakers more recent applications of our profession to national defense and homeland security.

Statistics in Defense and National Security: You Need More Than Statistics; You Need the Right Statistics

✱Nancy L. Spruill, Office of the Secretary of Defense, 3000 Defense Pentagon, OUSD(AT&L/ARA), Washington, DC 20301-3000, *Nancy.Spruill@osd.mil*; Ernest A. Seglie, DOT&E OSD ; Alyson G. Wilson, Los Alamos National Laboratory

Key Words: Defense, National Security, Testing, Reliability, Analysis, Data

The Defense and National Security communities depend heavily on two important aspects of statistics: data and analysis/reliability testing. Yet too often decisionmakers find themselves awash in data/statistics but find out too late that they do not have actionable information and/or statistical tools have not been correctly applied/better techniques exist. Statisticians are working to overcome these shortcomings by implementing the ASA's 2006 Task Force on Defense and Security, pursing new approaches by ASA's Section on Statistics in Defense and National Security, and responding to outreach by defense and national security professionals. This paper describes initiatives/examples where improvements resulted from involvement of statistics/statisticians, including a simple technique for cleaning up data, use of data mining, and reliability testing for an air to surface standoff missile.

Statistics in Defense and National Security: Past, Present, and an Eye to the Future

✱James R. Thompson, Rice University, 6100 Main Street, Dept. of Statistics, Houston, TX 77005, *thomp@rice.edu*

Key Words: beltway bandits, ombudsmen, black boxes, validation, simulation, data analysis

It should be a matter of concern that professional statisticians, in matters of defense, have largely lost the modeling, simulation and data analytical functions of our profession. To a large degree, contractors with ill-defined approaches, which are not subjected to validation from field data, have taken over these responsibilities. Statisticians are the de facto ombudsmen of scientific inference. Yet, since the McNamara Era, we have stood by and watched policies, combat models, and black boxes thrust upon our military without any vetting of their underlying assumptions or stressing of their predictive utilities by actual data. We statisticians need to stop going along with the system and start making waves.

5 Recent Bayesian Modeling and Computation Advances Published in *JCGS* ●

JCGS-Journal of Computational and Graphical Statistics, Section on Bayesian Statistical Science
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Bayesian Model-Based Clustering Procedures

✱John W. Lau, University of the Witwatersrand, School of Statistics and Actuarial Science, Private Bag 3, WITS 2050, Johannesburg, South Africa, South Africa, *john.w.lau@googlemail.com*; Peter J. Green, University of Bristol

Key Words: Dirichlet process, Hierarchical clustering, Loss functions, Stochastic search

We present a broad class of models for nonparametric Bayesian clustering of parametric gene expression profiles. Under a surprisingly wide range of prior assumptions, 'incremental' Gibbs samplers for the unknown partition are available for posterior simulation. The main focus of the talk is on optimal clustering under pairwise coincidence loss functions, and we present a novel heuristic algorithm for the resulting optimization problem.

Calibrating Environmental Engineering Models

✱David Ruppert, Cornell University, 225 Rhodes Hall, Ithaca, NY 14850, *dr24@cornell.edu*

Key Words: MCMC, Radial basis function, Data transformation

We present a Bayesian approach to model calibration when evaluation of the model is computationally expensive. Calibration is a nonlinear regression problem: given data Y corresponding to the regression model f(b), find plausible values of b. As an intermediate step, Y and f are embedded into a statistical model allowing transformation and dependence. Typically, this problem is solved by sampling from the posterior distribution of b given Y using MCMC. To reduce computational cost, we limit evaluation of f to small number of points chosen on a high posterior density region found by optimization. Then, we approximate the log-posterior using radial basis functions and use the resulting cheap-to-evaluate surface in MCMC. Our method can produce results similar to those when the true "expensive" posterior is sampled by MCMC while reducing computational costs.

Improved Estimation of Normalizing Constants from MCMC Output

✱Perry de Valpine, University of California, Berkeley, Environmental Science, Policy & Management, 137 Mulford Hall #3114, Berkeley, CA 94720-3114, *pdevalpine@nature.berkeley.edu*

Key Words: MCMC, normalizing constant, state-space, hierarchical model, mixed model, Bayes factor

Normalizing constants of conditional distributions are central to model comparisons. They include likelihoods of mixture models, such as hierarchical models and state-space time-series models, and Bayesian marginal likelihoods. I give three improvements to the bridge-sampling approach of Chib and Jeliazkov (2001) and Mira and Nicholls (2004). A different proposal density is used for the normalizing constant estimation than for the MCMC run. A ratio of effective sample sizes is included in the bridge function. And a Moving Block Bootstrap estimate of the Monte Carlo variance of the normalizing constant estimate is used to optimize algorithm parameters. I also show the full conditional is the optimal proposal density for estimating the normalizing constant. Comparisons from likelihood estimation for a state-space model show order-of-magnitude improvements in some cases.

6 Advances in Functional Data Analysis ●

Section on Statistical Computing, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, IMS

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Functional Regression in a Marketing Context

✱Gareth M. James, University of Southern California, Bridge Hall, 401R, Information and Operations Management Department, Los Angeles, CA 90089-0809, *gareth@usc.edu*

Key Words: Functional Regression, Penetration Data, Functional Clustering, Functional PCA, Marketing Data

Since Bass's landmark paper of 1969 parametric approaches such as the Bass model have played an influential role in prediction methods within Marketing. Nonparametric techniques such as FDA are relatively unknown in the Marketing Literature. In this talk we apply a Functional Regression approach to penetration curves corresponding to 760 product-country combinations. Our aim is to predict future penetration levels, time of product takeoff as well as time and level of peak sales. We compare the nonparametric functional approach to the Bass model in addition to several other approaches and demonstrate that considerable improvements in prediction accuracy are possible. In the process of producing our Functional Regression model we also illustrate that Functional Principal Components and Functional Clustering methods provide useful insights into the behavior of the penetration curves.

Forecasting with Functional Data in E-commerce

✱Wolfgang Jank, University of Maryland, 4322 Van Munching Hall, College Park, MD 20742, *wjank@rhsmith.umd.edu*; Natasha Foutz, University of Virginia; Gareth M. James, University of Southern California

Key Words: forecasting, principal component analysis, partial information, shape analysis, electronic commerce, prediction market

In this talk, we discuss several approaches of forecasting with functional data. We distinguish between short-term and long-term forecasting scenarios and we outline difficulties (and some solutions) for each scenario. In particular, we consider the problem of only partially observed information and how functional data analysis can be used to overcome it. We illustrate our methods in the context of electronic commerce where we study the problem of forecasting innovation success early and accurately. To that end, we use data from the Hollywood Stock Exchange, which is a prediction market used to forecast the success of Hollywood movies.

Modeling Sparse Generalized Longitudinal Observations with Latent Gaussian Processes

✱Fang Yao, University of Toronto, Department of Statistics, 100 St. George Street, Toronto, ON M5S3G3 Toronto, *fyao@utstat.toronto.edu*; Hans G. Müller, University of California, Davis; Peter G. Hall, The University of Melbourne

Key Words: Binomial Data, Functional Principal Component, Functional Data Analysis, Generalized Repeated Measurements, Prediction, Smoothing

In longitudinal data analysis one frequently encounters non-Gaussian data that are repeatedly collected for a sample of individuals over time. These observations could be binomial, Poisson or continuous. The timings of the repeated measurements are often sparse and irregular. We introduce a latent Gaussian process model for such data, establishing a connection to functional data analysis. We develop functional principal component (FPC) analysis for this situation and demonstrate the prediction of individual trajectories from sparse observations. This method can handle missing data and leads to predictions of the FPC scores. We illustrate these nonparametric methods with longitudinal data on primary biliary cirrhosis and show in simulations that they are competitive in competitive in comparisons with Generalized Estimating Equations and Generalized Linear Mixed Models (GLMM).

Structural Tests in Regression on Functional Variables

✱Laurent Delsol, Institut de Mathématiques de Toulouse, 118 route de Narbonne, Toulouse, International 31062 France, *delsol@cict.fr*

Key Words: structural test, functional variable, regression, nonparametric

Many papers deal with structural testing procedures in multivariate regression. More recently, various estimators have been proposed for regression models involving functional explanatory variables. Thanks to these new estimators, we propose a theoretical framework for structural testing procedures adapted to functional regression. The procedures introduced are innovative and make the link between former works on functional regression and others on structural testing procedures in multivariate regression. We prove asymptotic properties of the level and the power of our procedures under general assumptions that cover a large scope of possible applications: tests for no effect, linearity, dimension reduction, etc. We propose a bootstrap procedure and study its properties on simulations. Finally, we apply various structural testing procedures on a particular functional data set.

7 Bridging Clinical Trials ●

WNAR, Biopharmaceutical Section, Section on Health Policy Statistics, Biometrics Section

Sunday, August 3, 2:00 p.m.–3:50 p.m.

A Bayesian Approach for Local Bridging Evaluation and Sample Size Determination in a Multiregional Trial

Yi-Hau Chen, Institute of Statistical Science, Academia Sinica; Ya-Chi Wu, Center for Drug Evaluation, Taiwan; ✱Mey Wang, Center for Drug Evaluation, Taiwan, 1F, No. 15-1, Sec. 1, Hangjou S. Rd., Taipei, International 100 Taiwan, *mwang@cde.org.tw*

Key Words: multi-regional trial, bridging evaluation, Bayesian approach, sample size determination

A multiregional trial, conducted in more than one region under a common protocol, has been proposed with the desire to allow for global drug development and registration. Further, a multiregional trial allows for bridging evaluation so that global information can be utilized to support the registration of the drug at a given local region. From a statistical point of view, such an "information sharing" purpose can be nicely achieved by appealing to suitable Bayesian analysis. We propose a Bayesian approach for evaluating the drug effectiveness at a local region in a multiregional trial, where the posterior probability that the drug is effective at the local region is used, with the prior information obtained from the global data. A corresponding Bayesian approach for determination of the required sample size for evaluating the drug effectiveness in a local region is also proposed.

Statistical Issues in Multiregional Clinical Trials

✱Yoko Tanaka, Eli Lilly and Company, Lilly Corporate Center, DC6161, Indianapolis, IN 46285, *yokot@lilly.com*

Multiregional clinical trials are commonly conducted for registration of a new drug in various countries. The advantages of multiregional clinical trials may be the overall efficiency, including the speed, cost reduction, and better generalization of the results. However, it may not always be clear how to design or evaluate such multiregional clinical trial data, nor what evidence would have to be shown to the regulatory agency in a country or region. Particular issues to be considered include differences in ethnicity, medical practice, patient population, and dose regimen; number of subjects needed in each country or region; method of analysis of "by-country" or "by-region" data; and interpretation of overall and "by-country" or "by-region" data. Although there may not be a gold standard statistical approach to these issues, there are a number one can adopt.

Some Considerations in Design and Analysis of Multiregional Clinical Trials

✱Yoichi Ii, Pfizer Japan, Inc, Shinjuku Bunka Quint Bldg., 3-22-7, Yoyogi, Shibuya-ku, Tokyo, International 151-8589 Japan, *yoichi.ii@pfizer.com*

Key Words: multi-regional clinical trials, ICH-E5, global drug development

Increasing desire for simultaneous global drug development has motivated industry and regulatory agencies to update ICH-E5, specifically addressing multiregional trials. Japanese agency has also published a local guidance on the topic (2007). Yet, much remain unclear. Empirical Bayes approaches allow borrowing of strength, but relying too much on information from other regions may be an issue for a small region or when there is an apparent heterogeneity. Kawai, et al. (2007) gives a rationale for a minimum sample size for the smallest region, which could cap the influence of other regions. Furthering this idea, how could one borrow information, for example, when there are intrinsic and extrinsic factors that could lead to differences in treatment effect across regions. We will discuss how such factors could be incorporated into the design and analysis of a multi-regional trial.

8 Statistics in the Colorado News ●▲

ASA, Colorado-Wyoming Chapter, WNAR

Sunday, August 3, 2:00 p.m.–3:50 p.m.

A Modeling Framework for Evaluation and Comparison of Mp-/Id-Nat Trigger Strategies for West Nile Virus Screening

✱Brad Biggerstaff, Centers for Disease Control and Prevention, 3150 Rampart Road, Fort Collins, CO 80521, *BBiggerstaff@cdc.gov*; Lyle R. Petersen, Centers for Disease Control and Prevention

Key Words: resampling, West Nile virus, decision analysis

Screening of the U.S. blood supply for West Nile virus (WNV) began in 2003, shortly after transfusion-associated WNV transmission was first documented and recognition that the risk of transmission was high during seasonal outbreaks experienced in the US. Because of the seasonal nature of WNV transmission, which is a mosquito-bird-mosquito cycle in which humans intervene, during periods of low natural transmission, donations are pooled for aggregate testing, while during periods of high natural transmission, donations are tested individually. A statistical framework for evaluation and comparison of different strategies for determining when to initiate individual screening and when to revert to pooled screening is presented. This framework is used to evaluate various strategies of potential practical use under different collection center size assumptions and different epidemic conditions.

Nanostatistics: Numeracy for Trace-Level Environmental Contaminants

✱Dennis Helsel, Practical Stats, 9278 S. Lark Sparrow Dr., Littleton, CO 80126-5232, *dhelsel@practicalstats.com*

Key Words: water resources, nondetect, water quality, environmental statistics, training

Today's environmental chemicals of interest include complex mixtures of pharmaceuticals (antibiotics, antidepressants), personal care additives and synthetic hormones at nanogram per liter levels in rivers and drinking waters. Their effects include subtle and cumulative disruptions in the brain or reproductive cycles. Environmental scientists today need statistical training relevant to their current work. Their needs have moved beyond computing n, assuming a normal distribution, and confidence intervals. They require better training in robust statistics, censored data techniques, permutation tests, Bayesian methods and experimental design for multivariate contaminants. Statisticians must work alongside environmental scientists to use and develop new tools for interpreting the results of complex studies of trace-level contaminants.

Graduation Rates and Dropout Rates in Denver Public Schools

✱Alan Davis, University of Colorado Denver, School of Education & Human Development, Campus Box 106, P.O. Box 173364, Denver, CO 80217-3364, *Alan.Davis@cudenver.edu*

Key Words: education, Colorado, graduation rates

In the spring of 2005, the Rocky Mountain News ran a week-long series of articles on dropout and graduation rates in the Denver Public Schools. The series was based on the first comprehensive cohort analysis of high school students in Denver, and computed a graduation rate for minority high school students of 44%, far lower than the rate reported previously by the district to the Colorado Department of Education (CDE). Subsequently, CDE mandated cohort-based reporting of graduation rates. Complex issues remain in the computation of graduation and dropout rates, as non-negligible numbers of students transfer in and out of high schools, are retained in grade or remain in school for a fifth or sixth year before graduating, or leave school involuntarily. This presentation compares the results of precise panel tracking systems with simpler methods of estimation based on cohort membership over time.

9 Multiplicities in Statistical Analysis ▲

IMS, Biopharmaceutical Section, Section on Bayesian Statistical Science

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Bayesian Adjustment for Multiplicity

✱James Berger, Duke University/SAMSI, Statistical and Applied Mathematical Sciences Institute, P.O. Box 14006, Research Triangle Park, NC 27709-4006, *berger@stat.duke.edu*; James G. Scott, Duke University

Key Words: Bayesian, multiplicity, multiple testing, empirical Bayes

Issues of multiplicity in testing are increasingly being encountered in a wide range of disciplines, as the growing complexity of data allows for consideration of a multitude of possible hypotheses (e.g., does gene xyz affect condition abc). Failure to properly adjust for multiplicities is being blamed for the apparently increasing lack of reproducibility in science. The first purpose of this presentation is to review the way Bayesian analysis adjusts for multiplicity and, in particular, to differentiate this adjustment from the Bayesian Ockham's razor effect. The second purpose is to consider the empirical Bayesian approach to multiplicity adjustment, and show how it can significantly differ from the Bayesian adjustment.

Bayesian Model Robustness: Novel Approaches That Avoid Multiplicities

✱Kenneth Rice, University of Washington, Box 357232, F-600 Health Sciences Building, Dept of Biostatistics, Seattle, WA 98195-7232, *kenrice@u.washington.edu*

Key Words: Bayesian methods, Model robustness, Estimating Equations

In the traditional Bayesian paradigm, the model provides the likelihood, and the likelihood is the essential link between prior and posterior. Given belief in the specified model, this logic is compelling. However, when a 'good' model is implausible, the logic is less satisfying; model-robustness is not a natural property of traditional Bayesian inference. Current Bayesian approaches to robustness face substantial challenges; model selection induces a large multiple-comparisons problem, and model-averaging requires priors on the space of putative models. This talk describes a novel, decision-theoretic approach to model-robustness. It is fully Bayesian, but instead of the traditional emphasis on picking a model, we instead concentrate on the quantity to be estimated. No model selection or averaging is required. Connections will be made to GEE, notably the 'sandwich' covariance estimate.

Bayesian and Frequentist Multiple Testing in Cancer Genome Sequencing

✱Giovanni Parmigiani, Johns Hopkins University, 550 North Broadway, suite 1103, baltimore, MD 21205, *gp@jhu.edu*; Simina Boca, Johns Hopkins Bloomberg School of Public Health

Key Words: False Discovery Rates, Empirical Bayes, Hierarchical Models, Genomics, Cancer

The purpose of cancer genome sequencing studies is to determine the nature and types of alterations present in a typical cancer and to discover genes mutated at high frequencies. In the first of these studies to investigate the whole genome, about 20 thousand genes were sequenced, using a two stage design where all genes were sequenced on a "discovery" set of samples, and then those in which at least one alteration was found were sequenced in an additional "validation" sample. The two-stage sampling, the rarity of mutations, the varied size and composition of genes, all contribute to generating an interesting and unusual testing ground for comparing Bayesian, empirical Bayesian and frequentist approaches to multiple testing. In this presentation I will describe and compare our implementations of these three approaches using both real and simulated data.

10 Training TAs To Teach in Graduate School and Beyond ▲

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Training TAs To Teach in Graduate School and Beyond

✱Jessica Utts, University of California, Irvine, Irvine, CA 92697, *jmutts@ucdavis.edu*; ✱Candace Berrett, The Ohio State University, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *cberrett@stat.osu.edu*; ✱Roger Woodard, North Carolina State University, Department of Statistics, Campus Box 8203, 2501 Founders Drive, Raleigh, NC 27695-8203, *woodard@stat.ncsu.edu*; ✱Deborah Rumsey, The Ohio State University, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *rumsey@stat.osu.edu*; ✱W. Robert Stephenson, Iowa State University, 327 Snedecor Hall, Department of Statistics, Ames, IA 50011-1210, *wrstephe@iastate.edu*

Key Words: statistics education, TA training, teaching

This session will discuss graduate teaching associate (TA) training from specific training programs around the country, including ideas from the TAs themselves. All of us who train TAs want our graduate students to have the best training they can get for their teaching during graduate school and beyond. Since many TAs do not have the opportunity to take education courses, TA training programs are important for the future of the discipline.

11 Survey Respondent Incentives: Research and Practice

Section on Government Statistics, Social Statistics Section, Section on Survey Research Methods

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Survey Respondent Incentives: Research and Practice

✱Robert Groves, The University of Michigan, Ann Arbor, MI 48106-1248, *BGroves@isr.umich.edu*; ✱Brian Harris-Kojetin, U.S. Office of Management and Budget, 20503, *bharrisk@omb.eop.gov*; ✱Richard Kulka, Abt Associates, Inc., 20814, *Richard_Kulka@abtassoc.com*; ✱Clyde Tucker, Bureau of Labor Statistics, RM #1950, Washinton, DC 20212-0001, *Tucker.Clyde@bls.gov*; ✱Geraldine Mooney, Mathematica Policy Research, 600 Alexander Park, Princeton, NJ 08540, *gmooney@mathematica-mpr.com*

Key Words: incentives, response rates

Incentive payments to survey respondents have been used extensively for many years, and considerable research evidence supports their effectiveness in improving cooperation and the speed and quality of response. It has become increasingly difficult to achieve high response in many key surveys, and the use of incentives has become increasingly widespread. In March 2008, a one day seminar on survey respondent incentives brought together survey practitioners and methodologists to exchange views and practices related to providing incentives to survey respondents to improve response rates. This invited panel of survey researchers consists of key participants in the seminar who will consolidate and synthesize the main issues and key themes that emerged from the seminar and offer their perspectives on current and best practices and suggest future directions for implementation and research.

12 Generalized Linear Mixed Models with Applications in Biometry ●▲

Section on Statistical Consulting, WNAR, Biopharmaceutical Section

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Modeling Mosquito Abundance in Thailand Using GLMMs

✱Linda J. Young, University of Florida, Department of Statistics, University of Florida, Gainesville, FL 32611-0339, *LJYoung@ufl.edu*; Robert H. Zimmerman, University of Florida

Key Words: radial smoothing, genearlized linear mixed model, over-dispersed, count data

Malaria is transmitted by mosquitoes. The abundance of mosquitoes is affected by the availability of breeding sites which, in turn, depends on rainfall. In Thailand, rainfall is cyclical with peak rainfall generally occurring in April and minimum rainfall being observed in September or October. For three years, monthly data on the abundance of mosquitoes was collected using human bait in three Thai villages that are relatively close to each other. In this work, a generalized linear mixed model (GLMM) analysis was conducted to determine whether there is a difference in the abundance of mosquitoes among villages after adjusting for differences in rainfall. Rainfall within each village was modeled with splines, and radial smoothing was used to account for the temporal correlation structure in rainfall. The analysis and statistical issues will be discussed in detail.

Analysis of Hatch Data in Poultry Science: Moving from the Arcsine Square Root Transformation to Generalized Linear Mixed Modeling with SAS PROC GLIMMIX

✱Kevin S. McCarter, Louisiana State University, 161 Agricultural Administration Bldg, Baton Rouge, LB 70803-5606, *mccarter@lsu.edu*; Yan Qi, Louisiana State University

Key Words: logistic regression, generalized linear mixed model, GLIMMIX, poultry science

Researchers in poultry science often study factors affecting hatchability. The response variable is typically the number of chicks that hatch, which in most designs varies according to a binomial distribution. Analysis of this type of data has usually been accomplished using software such as SAS PROC GLM, after applying the arcsine square root transformation to stabilize the asymptotic variance. With the development of generalized linear mixed modeling and its recent availability in software such as SAS PROC GLIMMIX, binomial data from experiments involving random effects can now be modeled using software designed specifically for that purpose, eliminating the need for the arcsine square root transformation. We compare the two procedures through simulation and through analysis of a hatch dataset. We recommend the use of PROC GLIMMIX when modeling proportions with random effects.

Modeling the Regeneration of Grass in Golfing Divots Using GLMMS

✱Bruce Craig, Purdue University, West Lafayette, IN 47907, *bacraig@stat.purdue.edu*

Key Words: GLMMs, longitudinal study

Divot marks are a major concern to golf course managers because they hurt the quality and appearance of a golf course. This is especially true in the hot summer months because the grass does not regenerate quickly. A turf grass researcher at Purdue performed a longitudinal split-plot experiment to study the effects of irrigation, grass cultivar, and reseeding with and without fertilizer. For the first 14 weeks of the study, a machine was used to take a standard divot out of each plot. Once a divot was created, the percent grass cover in the divot was measured weekly until the end of the experiment. In this talk, I will discuss the use of generalized linear and nonlinear mixed models to analyze these data.

Spatial-Temporal Generalized Linear Mixed Models

✱Yanbing Zheng, University of Kentucky, 873 Patterson Office Tower, Department of Statistics, University of Kentucky, Lexington, KY 40506, *yanbing.zheng@uky.edu*; Jun Zhu, University of Wisconsin-Madison; Brian Aukema, Canadian Forest Service/University of Northern British Columbia

Key Words: generalized linear mixed model

Generalized linear mixed models provide a general framework for modeling spatial and/or temporal dependence via the specification of spatial-temporal random effects, but are computationally challenging for large data size. In this talk, we explore different ways of modeling the spatial and temporal random effects and devise computationally feasible algorithms for Bayesian inference. For illustration, we analyze real data examples in a study of mountain pine beetle outbreaks in western Canada.

Generalized Linear Mixed Models and Resource Selection Analyses

✱Kenneth W. Horton, Air Force Academy, 5921 Fossil Dr., Colorado Springs, CO 80923, *kwhorton1@gmail.com*; J. Richard Alldredge, Washington State University

Key Words: Spatial Modeling, Habitat Selection, GLMM

There are many statistical methods used to compare relative use and availability of resources (e.g., wildlife habitats). Most of these methods assume independence of resource selection among animals. In reality, this assumption is violated if animals display antisocial or gregarious behavior resulting in dependency among observations. This dependency can have a great influence on inference for resource selection data. We compare application of generalized linear mixed models for analysis of spatially dependent resource selection data with competing methods and make recommendations for use.

13 Model-Based Inference and Microarray Applications ●

Biometrics Section, Biopharmaceutical Section, WNAR

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Modeling the Distribution of P-Values in High-Dimensional Testing Applications

✱Gary L. Gadbury, Kansas State University, Department of Statistics, Dickens Hall, Manhattan, KS 66506, *gadbury@ksu.edu*

Key Words: FDR, correlation, microarray, plasmode

Many statistical methods model P-values from multiple tests of hypotheses to analyze data from high-dimensional experiments. These methods generally rely on uniformly distributed P-values under a null hypothesis. A null P-value will not be uniform if an incorrect reference distribution was used, and correlation among P-values from multiple tests can yield a distribution that is highly variable. Both of these issues can contribute to misleading "signals" in a distribution of P-values, and methods that model it may produce misleading results. Such signals can be present in an actual distribution of P-values, and how to incorporate them into simulated data when evaluating a method's performance is challenging. This challenge is discussed in the context of some data, a data derived simulation procedure is proposed, and the performances of some methods for estimating FDR are compared.

Criterion-Based Model Selection and Model Selection Philosophy

✱Kenneth P. Burnham, U.S. Geological Survey, and CSU, 201 Wagar Bldg., Colorado State University, Fort Collins, CO 80523-1484, *kenb@lamar.colostate.edu*

Key Words: model selection, model uncertainty, AIC, BIC, true model

Assessment of model selection uncertainly is easily done when model selection is criterion- and likelihood-based. Two common criteria are 1), select the true model (e.g., BIC) or 2), select the best predictive model (e.g., AIC). I elaborate on these ideas, in particular the deep basis for AIC (Kullback-Leibler information loss), and question the meaning of a "true" model. Under either criterion one has posterior model probabilities which are interpretable in terms of the selection criterion. This interpretation will be specified. Also relevant is that the target model for AIC varies with sample size, but if "true" model is the target criterion, the target model is independent of sample size. In reality, selection performance is dependent on sample size. One conclusion: the true model ("truth") as a model selection criterion is quite problematic.

Toward a Characterization of Gene Expression Variation in Tumors

✱Joseph Lucas, Duke Institute for Genome Science and Policy, NC 27514, *joe@stat.duke.edu*

Key Words: gene expression, microarrays, latent factor, factor regression, variable selection, cancer

We present a unified modeling framework for the identification of gene expression variation in human breast cancer. Sparse factor models are used to build gene expression factors from groups of co-expressed genes. These factors provide quantitative assessment of the activity of various biological pathways in cancer. Where possible, we identify the biological origin of these factors. Subsequently, we show that some are conserved across multiple tissue types, and that the level of activity of various pathways can be used in risk stratification and for the identification of drug susceptibility. Within the modeling framework, we are able to perform predictive analysis even across studies which show significant between study bias.

MCMC Inference for a Model with Sampling Bias: An Illustration Using SAGE Data

✱Russell L. Zaretzki, University of Tennessee, Dept of Statistics, Operations and Management Science, 328 Stokely Management Center, Knoxville, TN 37996-0532, *rzaretzk@utk.edu*; Mike Gilchrist, University of Tennessee; William M. Briggs, Methodist Hospital; Artin Armagan, The University of Tennessee

Key Words: Gibbs Sampling, SAGE, Gene Expression, Biased Sampling

This talk explores Bayesian inference for a biased sampling model in situations where the population of interest cannot be sampled directly, but rather through an indirect and inherently biased method. Observations are viewed as being the result of a multinomial sampling process from a tagged population which is, in turn, a biased sample from the original population of interest. This paper presents several Gibbs Sampling techniques which efficiently sample from the joint posterior distribution of a very large multinomial parameter vector. Samples from this method can be used to generate both joint and marginal posterior inferences. We apply it to a tagged population of messenger RNAs (mRNA) generated using a common high-throughput technique, Serial Analysis of Gene Expression (SAGE). Inferences for the mRNA expression levels in the yeast Saccharomyces cerevis are reported.

14 Study Design Issues in Medical Device Clinical Trials

Biopharmaceutical Section, WNAR, Biometrics Section

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Using Propensity Scoring Methods in Medical Device Studies That Include Both FDA Approval and CMS Reimbursement Endpoints

✱Dennis King, STATKING Consulting, Inc., 759 Wessel Drive, Unit 6, Fairfield, OH 45014, *statking@statkingconsulting.com*

Key Words: CMS Reimbursement, FDA Approval, Design of Medical Device Studies

Without reimbursement approval from CMS for a medical device, an FDA approved device is likely to be underutilized in the marketplace. Some authors have suggested the specification of endpoints and the collection of data to support both FDA approval and a favorable CMS reimbursement decision be integrated into clinical study protocols. Many pivotal trials for medical devices are observational studies in which the investigator has little or no control over the treatment assignment. Propensity score analysis methods have been used in the analysis of data from this type of clinical study to adjust the causal treatment effect for covariates. In this talk, the use of propensity score methods will be discussed in the context of FDA approval endpoints and CMS reimbursement endpoints.

Simulated Sample Size Required for a Geographical Region Relative to a Global Study in a Medical Device Trial

✱Jian Huang, Boston Scientific Corporation, 4100 Hamline Avenue North, South Town Square MS 9-315, St. Paul, MN 55112, *huangj4@bsci.com*; Haiying Lin, Boston Scientific Corporation; Yukiko Imai, Boston Scientific Corporation

Key Words: Global trials, Sample Size, Simulation, Superiority, Non-inferiority

The International Conference on Harmonisation (ICH - E5, 1998) provides the guidance on Ethnic Factors in the Acceptability of Foreign Clinical Data to minimize the duplication of clinical data and to allow for the extrapolation of those data to 'New regions.' For a global medical device trial, it is necessary to select the appropriate sample size for the regions. There is no well established method for region sample-size determination. In this presentation, the simulation method is proposed for evaluating the approaches recommended by Japan Pharmaceuticals and Medical Devices Agency (PMDA). A superiority and noninferiority design are discussed. Consistency across multiple regions is also examined.

Statistical Issues in Adaptive Design Trial

✱Yonghong Gao, U.S. Food and Drug Administration, 20850, *yonghong.gao@fda.hhs.gov*

Key Words: adaptive design, sample size, type I error rate

Early stopping for futility, early stopping for success, increasing sample size based on interim data and paying penalty in final analysis to control the overall type I error rate are important statistical elements in adaptive design trial. This presentation will look at these statistical issues and compare some plausible adaptive designs.

Adaptive Group-Sequential Design and Analysis for Correlated Binary Data on Repeated Venous Accessions

✱John C. Evans, Boston Scientific Corporation, 100 Boston Scientific Way, Marlboro, MA 01752-1242, *evansj1@bsci.com*; Brian P. Johnson, Boston Scientific Corporation; So Jung Imm, Boston Scientific Corporation

Key Words: Adaptive group-sequential design, Sample size re-estimation, Repeated measures, Correlated binary data

We present adaptive group-sequential designs for correlated binary repeated measures in a venous access study to compare proportions of complications on implantable ports. We estimate the correlation and average frequency of accessions at the interim analysis, and then use them to re-estimate the sample size for the final analysis. Whilst allowing for both the interim analysis to be done, and the correlation and frequency of accessions to be estimated, we ensure preservation of the overall significance level. We compare sample sizes and interim and final analyses under the usual assumption of independence of 13 accessions, with those obtained when an unknown positive correlation is assumed to be 0.1 among 13 accessions, and with those obtained when the correlation and average frequency of accessions are both estimated at the interim analysis.

Sample Size Simulation for a Post-Market Surveillance Study

✱Hongyan Qiao, Medtronic, Inc., 8200 Coral Sea St. NE, Mounds View, MN 55112, *hongyan.qiao@medtronic.com*

Key Words: Sample Size, Survival Rate, Precision

The data in a post-market surveillance study can reveal unforeseen adverse events, the actual rate of anticipated adverse events, or other information necessary to protect the public health. One of the main objectives of post-market surveillance studies is to monitor the performance of market released products. In order to assure meaningful product survival estimates, an enrollment strategy needs to be established. Many factors are related to calculation of sample size for survival data: general expectation for products survival rate, precision at the desired confidence level, and the expected dropout rate for the study. MEDSTAT (Medtronic, Inc.) software was used to assess sample sizes for a variety of different situations. The results provide a very valuable resource for determining an appropriate enrollment strategy for studies with objectives involving survival estimates.

15 Time Series I: Seasonal Adjustment

Business and Economics Statistics Section
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Seasonal Adjustment in a Mass Production Environment

✱Gary C. Brown, Office for National Statistics, Cardiff Road, Newport, International NP10 8XG United Kingdom, *Gary.Brown@ons.gov.uk*; Duncan J. Elliott, Office for National Statistics

Key Words: Infrastructure, Investigations, R

A common problem in seasonal adjustment is the sheer number of series to be analyzed. Combine this with a manually intensive analysis process for each series, and resources can become stretched. This paper reports how the Office for National Statistics in the UK is automating its seasonal adjustment process in order to deal with a massive increase in the number of series to be analyzed (a fall-out of re-organization of National Accounts). Efficiency benefits are already being felt, in advance of the anticipated avalanche of new work, as a result of the 'proof of concept' stages of this project. The paper takes the audience from mapping the current process, through identifying potential areas for automation and initial automation methods, to finding a long-term solution for this perennial problem.

Modeling of BLS and Census Bureau Seasonal Time Series with Frequency-Specific Generalized Airline Models

✱David Findley, U.S. Census Bureau, Statistical Research Division, Washington, DC 20233, *david.f.findley@census.gov*; Thomas Evans, Bureau of Labor Statistics; Jean Palate, National Bank of Belgium; Richard Tiller, Bureau of Labor Statistics

Key Words: ARIMA models, seasonal adjustment, forecasting

We present results from a study of the applicability of the Frequency Specific Models (FSMs) of Aston, Findley, Wills and, Martin (2004) to a broad range of seasonal time series, from the Bureau of Labor Statistics and from the Census Bureau. These models are generalizations of the most widely used seasonal ARIMA model, the (0,1,1)(0,1,1) or "airline" model. They make possible improved modeling of series whose seasonal movements are dominated by frequency components with frequencies in a subset of the seasonal frequencies 1, 2, 3, 4, 5 and 6 cycles per year (for monthly data). The study was done with a versatile menu-driven program, named GenairNBB, developed by the National Bank of Belgium. The software has batch options for processing large numbers of series that will be demonstrated. It will be available from the web site of the bank.

Using Besov Spaces and Empirical Mode Decomposition for Seasonal Extraction in Nonstationary Time Series

Christopher Blakely, Statistical Research Division; ✱Christopher Blakely, Statistical Research Division, *christopher.blakely@census.gov*

Key Words: Seasonal Adjustment, Empirical Mode Decomposition, Besov Spaces, Wavelets

The purpose of this paper is to introduce an empirical analysis technique for seasonal extraction in nonstationary time series. The proposed method is a nonmodel based approach to signal extraction, which uses analysis techniques borrowed from harmonic analysis including certain wavelet characterizations and empirical mode decompositions of the time series. We give a detailed account of the seasonal extraction algorithm followed by numerical examples, which both verify and validate the methods accuracy and robustness.

European Guidelines on Seasonal Adjustment

✱Gian Luigi Mazzi, European Commission, 5 rue Alphonse Weicker, Luxembourg, International 2721 Luxembourg, *gianluigi.mazzi@ec.europa.eu*; Rosa Ruggeri Cannata, European Commission; Cristina Calizzani, European Commission

Key Words: seasonal adjustment, calendar adjustment, harmonisation, revisions analysis

This paper presents the new European guidelines on seasonal adjustment. The choice of producing a detailed set of guidelines instead of a very synthetic one, as it is the case for the US and Canada, is discussed. Due to a lack of harmonization of seasonal adjustment practices at European level, it was necessary to produce a detailed set of guidelines giving specific recommendations for each step of seasonal adjustment. The paper presents the main characteristics of the guidelines stressing their innovative aspects and structure; for each item of the guidelines theoretical founding for the proposed recommendations as well as practical implications are discussed. Examples of the application of the guidelines using both Tramo-Seats and X-12 ARIMA are also presented to illustrate the impact of guidelines' recommendations on current seasonal adjustment practices.

The Application of X-13A-S on Monetary and Financial Data

✱Fida Hussain, Bank of England, Room H04(B-C), Threadneedle Street, London, International EC2R 8AH United Kingdom, *fida.hussain@bankofengland.co.uk*

Key Words: X-12-ARIMA, X-13A-S, seasonal, adjustments

The Bank of England (Bank) produces and publishes a wide range of monetary and financial time series, often in seasonally adjusted form. The Bank uses an in house developed software system (MIDAS) which calls the X-12-ARIMA seasonal adjustment package to seasonally adjust data. The data are generally reported as end point stocks. The Bank then uses other information to adjust the stocks for revaluations and other volume changes. We explore whether X-12-ARIMA can automatically estimate these adjustments and compare them to the actual adjustments currently employed. We present an empirical assessment of how the X-12-ARIMA package and the new X-13A-S package perform on a selection of Money and Lending series. We consider future directions and issues for the production of seasonally adjusted estimates at the Bank.

16 Student Paper Competition: Bayesian Methods for Bioinformatics and Binary Data

Section on Bayesian Statistical Science, Section on Teaching Statistics in the Health Sciences, WNAR

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Bayesian Model Search for Genetic Association Studies

✱Melanie A. Wilson, Duke University, 5500 Fortunes Ridge, Apt 99A, Durrham, NC 27713, *melanie.wilson@duke.edu*; Edwin S. Iversen, Duke University ; Merlise Clyde, Duke University; Scott C. Schmidler, Duke University ; Joellen M. Schildkraut, Duke University

Key Words: AIC, Bayesian Model Averaging, BIC, Evolutionary Monte Carlo, Model Selection, SNPs

Technological advances in genotyping have facilitated association studies of increasing scale. Accompanying this is a growing demand for statistical methods that are simultaneously more flexible and better able to handle larger numbers of variables. Many of the current approaches do not cleanly address the issue of multiple comparisons, correlations among the genetic markers and choice of their parameterization. We describe a Bayesian model search technique applied to penalized logistic regression that searches the space of genetic markers, and over the genetic parameterizations of each, that addresses these issues. This technique allows one to estimate posterior probabilities of both global and marginal associations. We characterize the performance of this approach in a simulation study and illustrate its use in two candidate pathway analyses.

Bayesian QTL Mapping for Multiple Traits

✱Samprit Banerjee, The University of Alabama at Birmingham, 1665 University Blvd, RPHB Rm # 327, Birmingham, AL 35205, *samban@uab.edu*; Nengjun Yi, The University of Alabama at Birmingham

Key Words: Bayesian, Seemingly Unrelated Regression, MCMC, multiple traits, QTL mapping, pleiotropy

Most QTL mapping experiments typically collect phenotypic data on multiple correlated complex traits. However, there is a lack of a comprehensive genome-wide mapping strategy for correlated traits in the literature. We develop a Bayesian multiple QTL mapping methods for correlated traits using two multivariate models; one which assumes the same genetic model for all traits, the traditional multivariate model, and the other known as the Seemingly Unrelated Regression (SUR) model allows different genetic models for different traits. We develop computationally efficient Markov chain Monte Carlo (MCMC) algorithms for performing joint analysis. We conduct extensive simulation studies to assess the performance of the proposed methods and to compare with the conventional single trait model. Our methods have been implemented in the freely available package R/qtlbim (*www.qtlbim.org*).

Bayesian Analysis of Longitudinal Binary Data Using Multivariate Bridge and Other Random Effects Models

✱Souparno Ghosh, Texas A&M University, 415 College Main Street, Apt #6, College Station, TX 77840, *sghosh@stat.tamu.edu*; Bani K. Mallick, Texas A&M University; Stuart R. Lipsitz, Harvard Medical School; Debajyoti Sinha, Florida State University

Key Words: Correlated binary data, Multivariate Normal distribution, Prbability integral transform, Bridge distribution, Copula model

A feature of random effects logistic regression models for longitudinal binary data is that the marginal link function, when integrated over the distribution of the random effects, is not of the same form as that of the conditional link. Recently, Wang et al. (2003) proposed a shared random intercept model in the clustered binary data to overcome this problem. In this article, we propose an extension of their model to model clustered binary data, with separate, but correlated, random effects following multivariate bridge distributions. The proposed copula model allows the marginal correlation among the clustered binary outcomes to decline with increasing time lag while retaining the property of having matching conditional and marginal logit link functions. Our novel model and associated methodology have been illustrated with the analysis of a clustered binary data from an AIDS study.

A Robust Bayesian Approach in Binary Data for Clinical Trials

✱Jairo A. Fuquene, University of Puerto Rico, Av Universidad # 4 Torre Norte, San Juan, PR 00925 Puerto Rico, *jairo.a.fuquene@uprrp.edu*; Luis R. Pericchi, University of Puerto Rico, Rio Piedras

Key Words: Bayesian Robustness, Clinical Trials, Exponential family

In clinical trials, binary data are ubiquitous. The usual Bayesian approach for binary data is beta/binomial conjugate analysis or a normal approximation to the log-odds and a conjugate normal prior. However, in this conjugate analysis, the influence of the prior distribution could be dominant even when prior and data are in conflict. This lack of robustness with respect to the prior is arguably the perceived weakest feature of Bayes, as Bayesian clinical trial analysis should be robust to the prior. If there is conflict between current data and past data, the past should be discounted. The perceived weakness is not of Bayes, but of conjugate (i.e., linear) Bayes. Our proposal is an analysis based on the Cauchy prior for the natural parameter in exponential families.

Reconstructing Tumor-Wise Protein Expression in Tissue Microarray Studies Using a Bayesian Cell Mixture Model

✱Ronglai Shen, Memorial Sloan-Kettering Cancer Center, *shenr@mskcc.org*; Jeremy M.G. Taylor, The University of Michigan; Debashis Ghosh, The Pennsylvania State University

Key Words: Tissue microarray, zero-augmented Gamma distribution, hierarchical Bayes model

A Cell Mixture Model (CMM) is proposed to reconstruct complex tissue staining patterns in TMA experiments. The concept is to assemble the whole-tumor expression pattern by aggregating over the subpopulation of tissue cores. Each individual core is assumed to be a zero-augmented Gamma distribution to assimilate the nonstaining areas and the staining areas. A hierarchical Bayes model is imposed to borrow strength across cores and across tumors. A joint model is presented to link the CMM expression model with a survival model for censored failure time observations. The implementation involves imputation steps within each MCMC iteration and Monte Carlo integration technique.

17 New Developments in Rank-Based Nonparametric Methods ▲

Section on Nonparametric Statistics

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Confidence Intervals for a Finite Population Median Based on the Sign Test

✱Suzanne Dubnicka, Kansas State University, Department of Statistics, 101 Dickens Hall, Manhattan, KS 66056, *dubnicka@ksu.edu*

Key Words: Horvitz-Thompson estimator, norm-based inference, superpopulation model

The construction of a confidence interval for the median of a finite population under unequal probability sampling will be discussed. A model-assisted approach makes use of the L1-norm to motivate the estimating function, which is then used to develop a unified approach to inference that includes not only confidence intervals but hypothesis tests and point estimates. The resulting hypothesis test is analogous to the ordinary sign test. We rely on large sample theory in most cases, and we will discuss the use of the Hartley-Rao variance approximation for cases in which second-order inclusion probabilities are difficult to obtain. Confidence intervals under simple random sampling without replacement, stratified random sampling, and cluster sampling will be explicitly illustrated.

Asymptotic Properties of Rank Likelihoods for Copula Models

✱Xiaoyue Niu, University of Washington, University of Washington, Department of Statistics, Seattle, WA 98195, *xiaoyue@stat.washington.edu*; Peter Hoff, University of Washington

Key Words: rank likelihood, asymptotics, copula model

Copula models provide a means of estimating associations in multivariate data separately from their univariate marginal distributions. For continuous data, the rank likelihood can be used as a marginal likelihood since the ranks are a partially sufficient statistic for the association parameters whose distribution does not depend on the univariate margins. We study its asymptotic properties and show that it provides consistent and efficient estimation. In the discrete case, the ranks depend on the marginal distributions. We generalize the rank likelihood to a special likelihood which depends only on the association parameters. We study the asymptotic properties of this extended rank likelihood and show that it performs as well as knowing the true margins. Therefore, it is appropriate to use this extended rank likelihood for mixed continuous and discrete data with arbitrary margins.

A Robust Rank-Based Procedure for Incomplete Longitudinal Data

✱Robin Mogg, Merck Research Laboratories, PO Box 1000, UG1CD-44, North Wales, PA 19454-1009, *robin_mogg@merck.com*; Devan V. Mehrotra, Merck Research Laboratories; Yang Liu, Florida State University

Key Words: Longitudinal trials, Multiple Imputation, Rank-based methods, Missing data

In a typical longitudinal comparative clinical trial, for a variety of reasons, some subjects discontinue the study before the end of their planned follow-up period. Restricted maximum likelihood (REML) is commonly used to analyze incomplete longitudinal data. REML can be highly inefficient in the presence of non-normal data; as such, many turn to generalized estimating equations

(GEE) and weighted-GEE when normality is untenable. As an alternative, we propose methodology that combines multiple imputation of the missing values with non-parametric (i.e., rank-based) methods. We study Type I error, power, bias, mean square error, width and coverage of 95% confidence intervals to show the robustness and power advantages of our proposed method relative to REML and (W)GEE. A numerical example and simulation results are used to illustrate the methodology and support our conclusions.

Comparisons of Rank-Based Estimators for GEE Models

✱Joseph McKean, Western Michigan University, 49008, *mckean@wmich.edu*; Ash Abebe, Auburn University; John Kloke, Pomona College

Key Words: Diagnostics, Iterated reweighted least squares, Monte Carlo, Nonparametric, Wilcoxon

Several recently proposed rank-based estimators of the effects in a GEE model are discussed. One of these estimators is based on an iterated reweighted least squares estimator for general linear and nonlinear models and another is based on a generalized R estimator. For both, weights can be chosen to bound influence in both the response and factor spaces. Illustrative examples and results of small sample investigations of this estimator and several other R estimators (which have optimal properties for specific GEE designs), are presented. Where appropriate, robust variance component estimators and related diagnostic procedures are also discussed.

Rank-Based Estimation for GARCH Models

✱Beth Andrews, Northwestern University, 2006 Sheridan Road, Evanston, IL 60208, *bandrews@northwestern.edu*

Key Words: GARCH, time series, rank estimation

A rank-based technique is used to estimate the parameters of generalized autoregressive conditionally heteroscedastic (GARCH) time series models. The estimators minimize the sum of mean-corrected model residuals weighted by a function of residual rank. Rank estimators are, in general, robust and relatively efficient. We show this is true in the case of GARCH parameter estimation. The estimation technique is robust because the rank estimators of GARCH model parameters are consistent and asymptotically normal under mild conditions. Since the weight function can be chosen so that rank estimation has the same asymptotic efficiency as maximum likelihood estimation, the rank estimators are also relatively efficient. In addition, rank estimation dominates classical Gaussian quasi-maximum likelihood estimation with respect to both robustness and asymptotic efficiency.

18 Risk Analysis for Industry and the Environment

Section on Risk Analysis, Section on Quality and Productivity

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Quantifying Local Creation and Regional Transport Using a Hierarchical Space-Time Model of Ozone as a Function of Observed NOx, a Latent Voc Process, Emissions, and Meteorology

✱Amy J. Nail, North Carolina State University, Campus Box 8203, NCSU, Raleigh, NC 27695-8203, *ajnail@stat.ncsu.edu*; John F. Monahan, North Carolina State University; Jacqueline Hughes-Oliver, North Carolina State University

Key Words: ozone, space-time, VOC, NOx, hierarchical, transport

We explore the ability of a space-time model to decompose the 8-hour ozone concentration at a point in space and time into the parts attributable to local emissions and regional transport. We model ozone as created plus transported ozone plus an error term that has a seasonally varying spatial covariance. We express created ozone as a function of the observed NOx concentration, the latent VOC concentration, and temperature. We model transported ozone as a weighted average of the ozone observed at all sites on the previous day, where the weights are a function of wind speed and direction. The latent VOC process model has a mean trend that includes emissions, temperature, a workday indicator, and an error term with a seasonally varying spatial covariance. Using likelihood methods, we fit the model and obtain predictions for comparison with a withheld data set.

An Analysis of the Potential Impact of Various Ozone Regulatory Standards

✱Rosalba Ignaccolo, Universita' degli Studi di Torino/SAMSI, 19 T.W. Alexander Drive, Research Triangle Park, NC 27709, *ignaccolo@samsi.info*; Yongku Kim, Statistical and Applied Mathematical Sciences Institute; Bahjat Qaqish, The University of North Carolina at Chapel Hill; Michela Cameletti, Universita' degli Studi di Bergamo/SAMSI; Richard L. Smith, The University of North Carolina at Chapel Hill

Key Words: risk assessment, respiratory-related mortality, rollback, ozone

We present a risk assessment analysis of the potential effect that various regulatory standards for ozone may have on the incidence of respiratory-related mortality. The analysis uses roll-back functions as models for the potential effect of regulatory standards. The statistical methods are based on the hierarchical Bayesian models. The objective is to obtain estimates of the effects of various regulatory standards, estimates of their variability, and the effects of various modeling assumptions on those estimates.

Multivariate Generalized Linear ARMA Processes: An Application to Hurricane Activity

✱Evangelos Evangelou, University of North Carolina, Department of Statistics and Operations Research, Chapel Hill NC 27599-3260, NC 27599-3260, *vangelis@email.unc.edu*; Richard L. Smith, The University of North Carolina at Chapel Hill

Key Words: Generalized linear ARMA processes, Maximum likelihood estimation

In this paper we propose a multivariate framework for investigating the relationship between hurricane activity and global warming. Papers such as Saunders and Lea (Nature, 2008) find evidence of correlation between the number of US landfalling hurricanes and local sea surface temperatures. We propose a modelling strategy involving a bivariate process where one component is Poisson and the other is Gaussian. Since standard time series analysis shows significant auto-correlations, we use a multivariate generalized linear ARMA model. Our analysis can be viewed as an extension of the methodology by Davis, Dunsmuir and Streett (2003, 2005) to multiple dimensions. Our maximum likelihood analysis shows that a multivariate framework can be a powerful tool for simultaneously analyzing hurricane activity and global warming in the presence of correlation between the two.

Probabilistic Risk Analysis for ICT Industry

✱Jose A. Rubio, Universidad Rey Juan Carlos, Tulipan s/n, Mostoles, International 28933 Spain, *jose.rubio@urjc.es*; David Rios Insua, Universidad Rey Juan Carlos

Key Words: Risk Analysis, Hardware/Software Systems, ICT, Bayesian analysis, Decision Analysis

The ICT Industry has been extremely active in developing risk analysis frameworks for their activities. However, while they have been doing an excellent qualitative job, their quantitative approaches tend to lack rigor. In this presentation we shall place on firm statistical and decision analytic ground one of the most popular risk analysis tools use in ICT, emphasizing key modeling issues.

Seismic Risk Analysis

✱Mircea Grigoriu, Cornell University, 14853, *mdg12@cornell.edu*

Key Words: Bayesian analysis, activity matrix, seismology, Gaussian processes

A method is developed for calculating the reliability of a physical system in seismic environment, that is, the probability that the system state does not leave a safe set D during a specified time interval $[0,\tau]$. It is assumed that (1)~system properties are perfectly known, (2)~activity matrices, that is, histograms giving relative frequencies p_k of earthquakes with parameters in sets C_k, $k=1,\ldots,r$, partitioning the space of earthquake magnitude and site to source distance, can be used to characterize site seismicity, (3)~time arrivals of seismic events at a site can be described by a homogeneous Poisson process with intensity inferred from the corresponding site activity matrix, and (4)~ground accelerations during individual seismic events constitute samples of a Gaussian processes with mean 0 and C_k-dependent spectral densities.

19 Methodological Challenges Encountered at the U.S. Centers for Disease Control and Prevention in the Division of HIV/AIDS Prevention ●

Section on Statistics in Epidemiology, Section on Government Statistics, Social Statistics Section, WNAR, Biometrics Section

Sunday, August 3, 2:00 p.m.–3:50 p.m.

On Validating a Stepwise Variable Selection Algorithm with Another Stepwise Variable Selection Algorithm

✱Ryan E. Wiegand, Centers for Disease Control and Prevention, Mailstop E48, 1600 Clifton Road, NE, Atlanta, GA 30329, *fwk2@cdc.gov*

Key Words: Variable selection, stepwise, validation

The backward elimination, forward selection, and stepwise variable selection algorithms have been criticized in statistical studies, yet their use is still common in medical, social science and environmental literature. One such application of these methods is to use a stepwise algorithm to build a multivariable model and employ a different stepwise procedure to validate the model. This research attempts to characterize the circumstances for which this process is an advisable variable selection strategy. Data for linear and non-linear regression models were simulated varying the number and type of variables and correlation between those variables. The final models of the three stepwise variable selection algorithms were compared for each simulation and the proportions of agreement and agreement on the correct model will be presented.

Identifying and Interviewing Persons with a Recent HIV Diagnosis Not Yet Receiving Medical Care

✱Christopher H. Johnson, Centers for Disease Control and Prevention, 1600 Clifton Road NE, MS E-48, Atlanta, GA 30330, *cjohnson3@cdc.gov*

Key Words: surveillance, HIV, surveys, misclassification bias

CDC surveillance systems monitor new HIV cases and patients receiving care, but little is known about persons not in care. CDC is therefore piloting the Never In Care Project. Diagnoses reported to the HIV/AIDS Reporting System are compared with CD4 and viral load test results, reportable by law to state laboratory reporting systems. Adults reported as HIV + but with no further post-diagnosis test results or other evidence of care entry within 90 days of diagnosis are eligible for the study. Sites will identify eligible subjects and collect data through September 2009, collaborating with CDC on methods, data management, and quality control. Interview data from some 1200 surveyed participants will provide information on barriers to care. CDC will assess selection and misclassification bias due to incomplete and delayed reporting and evaluate system design and utility.

An Improved Procedure for Accounting for Reporting Delay

✱Ruiguang Song, Centers for Disease Control and Prevention, Atlanta, GA 30333, *rsong@cdc.gov*

Key Words: Adjustment weight, Bias, Reporting delay, Trend

Adjusting for reporting delay is important in trend analysis of HIV/AIDS incidence and prevalence. Limitations of the current approach for estimating reporting delay weights include its questionable assumptions that the reporting delay distribution has not changed in the past five years and that no delay exceeds five years. Its complicated formation of homogenous reporting delay groups is based on a reverse-time proportional hazards model that may combine population groups with different reporting delay patterns. A new weight generation procedure overcomes these problems using regression to estimate reporting delay probabilities, a parametric function to model delays longer than five years, and cluster analysis to identify homogenous reporting delay groups. Advantages of the new procedure are demonstrated with analytic and simulation results.

Use of Population-Based Data Bases To Estimate HIV/AIDS-Relative Survival

✱Xinjian Zhang, Centers for Disease Control and Prevention, Division of HIV/AIDS Prevention, 1600 Clifton Road, Atlanta, GA 30333, *xaz8@cdc.gov*; Ruiguang Song, Centers for Disease Control and Prevention; Kathleen McDavid Harrison, Centers for Disease Control and Prevention; Gengsheng (Jeff) Qin, Georgia State University

Key Words: Human Immunodeficiency Virus (HIV), Acquired Immunodeficiency Syndrome (Aids), Relative Survival

Relative survival after HIV or Aids diagnosis, the ratio of observed and expected survival probabilities can be estimated using routinely collected population-based data. Observed survival was calculated using data from the HIV/Aids reporting system (HARS) at the Centers for Disease Control and Prevention (CDC) which reflects the survival experience of all HIV and AIDS patients diagnosed in the United States. Expected survival was obtained from US Census data for the general population. A multivariate generalized linear model, using exact survival times, grouped covariate data, and Poisson error structure, fit the data well. The overall relative survival after AIDS diagnosis was 0.872 and 0.778 at 1 and 3 years (p<0.0001) respectively; after diagnosis of HIV infection, it was 0.932 and 0.882 at 1 and 3 years (p<0.0001) respectively.

20 Data-Driven Methods on Survey Data ●

Section on Survey Research Methods, Section on Nonparametric Statistics, Section on Government Statistics, SSC, Social Statistics Section

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Building Regression Trees on Survey Data

✱Daniell Toth, Bureau of Labor Statistics, Office of Survey Methods Research, 2 Massachusetts Avenue NE, Washington, DC 20212, *toth.daniell@bls.gov*; John L. Eltinge, Bureau of Labor Statistics

Key Words: CART, Machine Learning, Pruning, Stratified Samples, Regression Trees, Survey Weights

In this paper we explore the effects of a complex sample design on the estimation of a regression tree. To do this we present a simple function representation for a regression tree that more closely resembles standard regression. From this representation we form an estimating equation that allows us to incorporate survey weights in our estimation of the coefficients. In addition our method of pruning allows us to incorporate weights so we can consider the impact of survey weights on the tree pruning procedure. We compare the population estimates to a tree built without incorporating survey weights.

Tree-Based Methods To Model Dependent Data

✱Guillermo Mendez, Arizona State University, *guillermogmendez@yahoo.com*

Key Words: suvey sampling, clustered data, regression trees, mixed random forest, small area estimation

In contrast to linear mixed models, most supervised learning methods do not take into account whether the data have some clustered structure. Random forest (RF) is a popular technique used to model complex data sets because it tends to produce accurate predictions. We propose a modified RF algorithm called Mixed Random Forest (MRF) that can model mixed-effects data using regression trees to produce robust estimates of group means and random effects. We compare the results of the MRF algorithm with those from linear mixed models for the small area estimation (SAE) problem. Our method is shown to perform better in terms of mean squared prediction error when the underlying function is complex, such as conditionally linear, even though no model is pre-specified. Although we focus on SAE, our method can be used whenever the data are clustered as long as the number of clusters is large.

Nonparametric Density Estimation from Censored Data

✱Meghan S. O'Malley, Bureau of Labor Statistics, 2 Massachusetts Ave, NE, Washington, DC 20212, *meghan.s.omalley@gmail.com*

Key Words: non-parametric, density estimation, censored data, histogram

This paper explores the accuracy of a simple density estimator, based solely on one histogram, for a variety of possible distribution shapes and bin/sample size combinations, through a simulation study. The density estimator is a piecewise quadratic polynomial chosen to match histogram areas with boundary points initially at midpoints of adjacent histogram bars then improved for smoothness. Performance is measured by the Mean Integrated Squared Error of the density estimates themselves and the Mean Square Errors of the means and a few percentiles derived from the density estimates. To give insight into the performance of this density estimator in practice, the piecewise quadratic density estimator is applied to wage data from the Bureau of Labor Statistics and compared to kernel density estimates using corresponding point data.

Can Survey Bootstrap Replicates Be Used for Cross-Validation?

Geoff Rowe, Statistics Canada; ✱David Binder, Statistics Canada, 49 Bertona Street, Nepean, ON K2G 4G7 Canada, *dbinder49@hotmail.com*

Key Words: Complex surveys, Model prediction error, Health Survey, Regression

We propose an extension to bootstrap methods for evaluating regression models estimated with data from surveys with complex design. Such methods involve selection of replicate samples formed from simple random samples of clusters within strata. Selection is carried-out with replacement, so that about one third of clusters are inevitably left out of a given replicate sample. Our evaluation method exploits the excluded clusters, using them as cross-validation samples for assessment of a model's prediction error; while using the bootstrap samples to estimate the variance of regression coefficients. The paper is in two parts. Part I outlines design-based properties of the bootstrap/cross-validation. Part II illustrates our method with a comparative assessment of selected models of health dynamics using Statistics Canada's National Population Health Survey (1992-2004).

21 An Overview of K–16 Poster and Project Competitions ▲

Section on Statistical Education

Sunday, August 3, 2:00 p.m.–3:50 p.m.

An Overview of K–16 Poster and Project Competitions

✱Carl Lee, Central Michigan University, PE 109, Department of Mathematics, Mt. Pleasant, MI 48859, *carl.lee@cmich.edu*; ✱Jerry Moreno, John Carroll University, Department of Mathematics and Computer Science, University Hts., OH 44118, *moreno@jcu.edu*; ✱Megan Mocko, University of Florida, P.O. Box 118545, Gainesville, FL 32611, *mmeece@stat.ufl.edu*; ✱Neal Rogness, Grand Valley State University, Statistics Department, MAK 1133, Allendale, MI 49401, *rognessn@gvsu.edu*

Key Words: Poster, Project, Competition, Undergraduate, Primary, Secondary

The panel members have all been involved with various poster or project competitions available for K–16 students that are sponsored by ASA or CAUSE. The panelists will share overviews of and insights into the competitions. Suggestions will be made as to how one can become involved in the events, including how to organize a local competition and how to encourage students to participate in the competitions.

22 Statistical Methods for Networked and Streaming Data ●

Section on Physical and Engineering Sciences, Section on Quality and Productivity, IMS

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Statistical Network Comparison

✱Wei Liu, U.S. Food and Drug Administration, CDER/OTC, 10903 New Hampshire Avenue, Silver Spring, MD 20993, *liuding@verizon.net*; Wolfgang Polonik, University of California, Davis

Key Words: network scrambling, topological dissimilarity, spectral analysis, statistical comparison, normalized and standardized adjacency matrix, moments

Statistics is challenging in the study of dynamical random networks, because in general only a very small number of observed networks is available. We are interested in statistical inference of topological dissimilarities between networks. We are proposing an approach for developing tools for statistical inference of large networks. The key underlying idea is to assesses topological dissimilarities between networks indirectly through adding noise to their structures - also called scrambling. The level of scrambling needed in order to make the topologies of the scrambled networks statistically indistinguishable is used as a dissimilarity measure. Topologies are compared via the spectral distributions of their (normalized and standardized) adjacency matrices. The methodology is illustrated via simulation studies and via an application to real protein-interaction networks.

Time Series Synchronization and Detection Using the Hamming Distance

Carolyn B. Morgan, Hampton University; ✱Morris H. Morgan, Hampton University, Queen and Tyler Streets, Hampton, VA 23668, *morris.morgan@hamptonu.edu*; Clementina Alexander, Hampton University

Key Words: time series, bootstrapping, confidence estimates

With the growth of wireless technology, interest in digital communication applications has accelerated. A problem of interest is detecting synchronization levels of signals (time series) arising from divergent sources. In such systems, the primary focus is on the statistical characterization of binary signals that carry encoded data. The Hamming distance has been used successfully in many applications ranging from developing an error code correcting device to facial recognition applications. In the present context, the Hamming Distance serves as a metric for accessing signal synchronization levels and performing statistical inferences about information reliability. The Hamming distance approach that employs the count or difference between two time series is shown to be a quick and effective technique for discrimination. Bootstrapping is then used to obtain confidence estimates.

Anomaly Detection in Internet Traffic Data

✱Nathaniel Beagley, Pacific Northwest National Laboratoty, 902 Battelle Blvd., Richland, WA 99352, *Nathaniel.Beagley@pnl.gov*; Paul Whitney, Pacific Northwest National Laboratory; Christian Posse, Pacific Northwest National Laboratory; Amanda White, Pacific Northwest National Laboratory

Key Words: Cyber-security, Anomaly Detection

The quantity of internet traffic flowing in and out of an institution has reached a size where, even making use of automated alert tools, it is impossible for a cyber-security analyst to fully examine the complete data set. Methods are needed to identify interesting or "anomalous" cases to allow the analyst to focus their efforts. In our work, we run two different anomaly detection algorithms (a Holt-Winters algorithm designed for robustness, and a least squares algorithm designed for efficiency) on a collection of time sequential data derived from internet session data sensors. The set of anomalies occurring at a specific time point represents the signature for that time point which is then compared against a library of signatures generated from "known" anomalous events to decide if the time point is a candidate "unknown" anomalous event.

Analyzing Diffusion Tensor Imaging Data

✱Meagan Clement, Rho, Inc., 6330 Quadrangle Dr., Chapel Hill, NC 27517, *meagan_clement@rhoworld.com*; David Couper, The University of North Carolina at Chapel Hill; Steve Marron, The University of North Carolina at Chapel Hill; Keith E. Muller, University of Florida; Hongtu Zhu, The University of North Carolina at Chapel Hill

Key Words: Diffusion tensor imaging, diagnostic tools, density estimation

Diffusion tensor imaging (DTI) holds tremendous promise for improving our understanding of neural pathways, especially in the brain. Unfortunately, little has been done to define metrics or describe credible statistical methods for analyzing DTI data. Using a transformation of the fractional anisotropy (FA) values will lead to outcomes in statistical models that avoid the "curse of dimensionality" (from having far more variables than independent sampling units). It is of the utmost importance to have the diagnostic tools available to determine if the models are accurate for the region of interest. Issues of combining different tissue types or different regions of the brain in an analysis could lead to a bimodal distribution of the transformed FA values. If this occurs, new analysis approaches must be considered. Analysis approaches for any region will be introduced.

Recurrent Event Data Under Competing Risks

✱Ananda Sen, The University of Michigan, 3550 Rackham, 915 E Washington, Ann Arbor, MI 48103, *anandas@umich.edu*

Key Words: Competing risks, Recurrent event, Masking, NHPP

The focus of the talk is recurrent events for which the relevant data comprise successive event times for a recurrent phenomenon along with a event-count indicator. Situations in which individuals or systems in some population experience recurrent events are common in areas such as manufacturing, software debugging, risk analysis, and clinical trials. While the different application areas dealing with recurrent data enjoy certain commonalities, often the issues specific to an area warrant a nonstandard modification of the basic methodologies. In this talk, I shall discuss competing risks analysis in the context of recurrent failures that are subjected to multiple failure modes. Models and methodologies for analyzing a single as well as multiple clusters of recurrent events will be discussed. The situation dealing with missing or masked cause of failure will further be addressed.

Using Labeled Data To Evaluate Change Detectors in a Multivariate Streaming Environment

Werner Stuetzle, University of Washington; Donald B. Percival, University of Washington; ✱Caren Marzban, University of Washington, Applied Physics Laboratory, 1013 NE 40th Street, Box 355640, Seattle, WA 98105-6698, *marzban@stat.washington.edu*; Albert Kim, University of Washington

Key Words: change, detection, two-sample, anomaly, novelty

We consider the problem of assessing the relative performance of different change detectors in a multivariate streaming environment. The change detectors of interest in this paper are two-sample test statistics, and assess whether the distribution of most recently observed vectors differs significantly from that observed previously. This assessment is difficult in a streaming environment because of the violation of the assumption that the vectors are IID. We argue that given labeled data one can compare the relative performance of different two-sample test statistics without unduly strong assumptions. To facilitate the comparison of change detectors, we propose adaptations of the standard notions of a receiver operating characteristic (ROC) curve, and of bootstrapping, appropriate for a labeled streaming environment.

23 Linear Modeling Methods for Education Statistics and Other Social Statistics ●

Social Statistics Section, Section on Survey Research Methods

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Effect of Spanish Two-Way Immersion Program on Kindergarten Students' Attitudes: A Study of Treatment and Control Schools

✱Claudia B. Navarro Villarroel, Iowa State University, N131 Lagomarcino Hall, Iowa State University, Ames, IA 50011, *cnavarro@iastate.edu*; Mack Shelley, Iowa State University; Marcia Rosenbusch, Iowa State University; Holly Kaptain, Iowa State University

Key Words: Generalized multinomial linear model, Survey for children, Dual Language Education

A two-way immersion (TWI) program is a type of dual language education with a "balanced" population of native English speakers and native speakers of the target language, such as Spanish. The present study was conducted with kindergarten students after their first year of participation either in a Spanish TWI or in a traditional program. The study was conducted in two rural and two urban schools in a Midwestern state, with a total sample size of 131 students distributed unequally through the schools. The children were asked to answer a 14-question paper survey about their attitudes toward school and learning. A generalized multinomial linear model analysis revealed that TWI programs caused significantly more positive effects in the kindergarten students' attitude toward the Spanish language in comparison to their peers in control schools.

Matching Four Groups of Postsecondary Education Institutions Using Propensity Scores

✱Amang S. Sukasih, Mathematica Policy Research, Inc., 600 Maryland Ave., SW, Suite 550, Washington, DC 20024, *asukasih@mathematica-mpr.com*; Kirsten Barrett, Mathematica Policy Research, Inc.; Wendy Mansfield, Mathematica Policy Research, Inc.

Key Words: matching, multiple treatments, IPEDS, CCAMPIS, childcare

Propensity score matching is widely used to form comparable "treatment" and "control" groups in observational studies. When forming these groups, a logistic regression model is commonly used to estimate scalar propensity scores by which the two groups are being matched. If the number of groups is larger than two, several methods are available to construct the groups, including generalized propensity scores and scalar balancing scores. This paper presents an application of these methods in selecting postsecondary education institutions into four comparable groups defined by cross classifying institutions on whether they received Child Care Access Means Parents in School (CCAMPIS) grants and whether they have on-campus child-care centers. Moreover, we compare these matching results to those based on simple sequential binary propensity score matching.

Methods for Analyzing Agreement on Ordinal Ratings Between Self- and Peer-Evaluations in a Medical School Student Survey

✱Oliver Cass, Hennepin County Medical Center, 701 Park Ave, Department of Medicine, Minneapolis, MN 55415, *cassx001@umn.edu*; Ye-Ying Cen, Hennepin County Medical Center

Key Words: Pearson correlation coefficient, weighted Kappa coefficient, multinomial, logistic, ordinal

A total of 171 students at a medical school self-rated on 3 questions measuring professionalism, teamwork, and mastery of content with ordinal ratings from 1-5. Each student was also rated by 3 different peers. Pearson correlation coefficients between self-evaluation and the average of 3-peer evaluations and Kappa coefficients comparing the consensus of 3-peer evaluations and the self-evaluation were computed. A multinomial cumulative logistic model was applied. Both Pearson correlation coefficients and weighted Kappa coefficients showed a significant agreement between self- and peer- evaluations. The multinomial logistic regression model suggested peer-evaluation tends to be more generous than self-evaluation. For example, if self-evaluation was 3, there was a 70% probability that peer-evaluation would be 4.

Modeling of Probability of Infant Survival (Survival Rates)

✱Beau Jones, University of Cenral Arkansas, Department of Mathematics, 201 Donaghey Ave, Conway, AR 72035, *beauj@uca.edu*; Evgenia Rubinshtein, University of Central Arkansas

Key Words: infant survival rate, modeling, missing data, regression

Modeling of infant survival rate is important for many reasons such as prediction U.S. population, prediction of health-care and daycare facilities, manufacturers (infant food, cloth etc.). We use regression model to simulate infant survival rate. Different approaches are used to handle missing data. Simulations is done for five Southern States including Arkansas, Texas, Tennessee, Alabama, and Missouri. The results of the modeling for 7 years are compared.

Deal with Zero Response When It Is Important Measure but Noisy

✱Gideon D. Bahn, Loyola University Chicago, 1623 Coolidge Ave., Wheaton, IL 60187, *gideonbahn@yahoo.com*; Raymond Massenburg, University of Illinois at Chicago

Key Words: homicide, foreclosure, subprime lending, Logistic regression, Hausman-Test, multiple linear regression

When the outcome variable has many missing data, imputation has been a common method. But in the case of homicide, zero response is important information which should not be imputed in the analysis. This paper examines the relationship among homicides, foreclosures and demographic factors in Chicago. In order to treat the zero response two methods are employed. Firstly, the data is analyzed using Logistic Regression Modeling to compare the difference between with and without homicide. Secondly, Linear Regression Modeling was utilized to understand the relationships between the homicide rate and the foreclosure rate; and between subprime lending rate and homicide rate. The simultaneity is checked in the second model through the Hausman-Test. The results indicate that the foreclosure rate and the subprime lending rate are positively significant predictors of the homicide rate.

An Examination of the Relationships Between University Student Retention and High-School Coursework Among At-Risk Students Using Regression Analysis

✱Chong Ho Yu, Arizona State University, 3S89 Computing Commons, Tempe, AZ 85287-0101, *chonghoyu@gmail.com*; Angel Jannasch-Pennell, Arizona State University; Samuel Digangi, Arizona State University; Charles Kaprolet, Arizona State University

Key Words: exact logistic regression, retention, attrition, conditioning, odds ratio

The objective of this project is to examine how high-school coursework affects attrition of university students using regression analysis. Specifically, previous studies repeatedly show that minority and nonresident students are more likely to drop out from universities. High school GPA and SAT scores do not tell us how performance in different disciplines is related to college persistence. In this study, archival data sets from the data warehouse of a Southwestern university, which contain records of students' high school coursework, will be utilized. Since a high volume of missing data is expected, the final sample size may be small and data points may be sparse. To counteract the problem, exact logistic regression and other remedies will be employed.

24 Psychometrics and Latent Class Analysis

Section on Survey Research Methods, Section on Government Statistics

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Measuring Nonresponse Bias from Response Propensities Obtained from Respondents' Answers on Their Survey Participation Behavior

✱William Kalsbeek, Unversity of North Carolina, Survey Research Unit, UNC-CH, 730 Martin Luther King Jr, Chapel Hill, NC 27607, *bill_kalsbeek@unc.edu*; Robert Agans, University of North Carolina

Key Words: Nonresponse bias, Response propensity

The nonresponse bias of simple linear estimates can be measured directly if the response propensity for each survey respondent is known. In reality these propensities are unknown and must be estimated. Traditional estimation methods include the weighting class adjustment and response propensity modeling. This research explores the use of psychometric methods to estimate propensities from ancillary information plus respondents' answers to questions about their availability and willingness to participate in surveys more generally. Data from two recent telephone surveys conducted by the UNC Survey Research Unit are used to compare estimates of nonresponse bias computed from propensities obtained by this approach versus bias from propensities obtained by the traditional methods.

Reducing Error in Psychological Research: Averaging Data To Determine Factor Structure of the QMPR

✱Shea Gibbons, Brigham Young University, 1005 SWKT, Provo, UT 84602, *sheagibbons@gmail.com*; Robert Bubb, Brigham Young University; Bruce L. Brown, Brigham Young University

Key Words: Factor Analysis, Reliability, Averaging Data, Human Subject Data, QMPR

Human subject data in psychological research often contain a high level of error. Factor analytic data are no exception. Factor loading instability is common in single administration factor analytic research and results in poor interpretation of the factor pattern. However, Monte Carlo simulations have shown that averaging data across administrations reduces error, resulting in increased explained variance and stable factor patterns. In this study, the Questionnaire for the Measurement of Psychological Reactance (QMPR) was administered multiple times. The data were then averaged and analyzed using principle-components factor analysis with varimax rotation. The results of the study show a notable increase in explained variance and factor-pattern stability, supporting previous simulation findings. Averaging data across administrations is advocated to reduce error in human subject data.

Comparison of Students' Scores Between Public and Private Schools: Analysis of PISA Data

✱Victoria Landsman, The Hebrew University of Jerusalem, Department of Statistics, Mount Scopus, Jerusalem, 91905 Israel, *msvika@mscc.huji.ac.il*; Danny Pfeffermann, Hebrew University & University of Southampon

Key Words: Informative sampling, Instrumental variables, Probability weighted estimators, Propensity scores, Sample distribution

We compare students' scores between public and private schools in Ireland, using data collected for PISA. This survey employs a two stage sampling design of schools and pupils, but the choice of school by pupils is not under control. Hence, pupils of the two types of school can differ in unknown characteristics, and direct comparison of the sample scores may lead to erroneous conclusions. We compare the performance of commonly used approaches to deal with this problem, such as Hajek and Greg type estimators with weights equal the inverse propensity scores, instrumental and latent variables models and an alternative approach that we developed. This approach fits a model to the data by modeling the outcome distribution under ignorable assignment and the school selection probabilities, which are modeled as functions of the outcome and covariates, thus accounting for informative selection.

Issues Matter: Social Attributes, Predispositions, and Issues Affect Voting Choice

✱Robert B. Smith, Cytel Inc., 675 Massachusetts Ave., 3rd Floor, Cambridge, MA 02139-3309, *rsmithphd@comcast.net*

Key Words: campaign issues, political philosophy, health care reform, economic issues, environmental advocacy, candidate's character

Clarifying the relative impacts of economic and social issues, this study analyzes a 1992 election night survey assessing how social attributes, philosophical self-designation (liberal, centrist, conservative), party identification (Democrat, Independent, Republican), and the issues influenced the voters' choices. Residents of east/west coastal regions, women, paid workers, and first time voters leaned toward liberalism; ethnic minorities, older people, and poor people espoused a Democratic identification. Political philosophy had a direct effect on party identification, which had a very strong direct effect on vote. A three-class latent structure analysis produced a Left-Center-Right classification. The issues influenced the vote. The economic issue had the strongest effect on Clinton vote.Issue interactions suggest that environmental advocacy weakens the effects of negative campaigns.

A Brief History of Survey Classification Error Models

✱Marcus Berzofsky, RTI International, 3040 Cornwallis Dr., RTP, NC 27709, *berzofsky@rti.org*; Paul P. Biemer, RTI International/UNC-CH; William Kalsbeek, University of North Carolina

Key Words: Latent class analysis, Classification error, Hui-Walter Method, Census Bureau Model

Classification error analysis aims to measure misclassification rates and to correct estimates of proportions for misclassification. We will trace development of the key models developed for this type of analysis and compare and contrast them. Of particular interest are the latent class models developed by Lazarasfeld and Henry (1968), the Census Bureau Model developed by Hansen, Hurwitz, and Pritzker (1964) and the finite mixture probability models developed by Tennebein (1972), Hui and Walter (1980) and others. This paper will show how these three primary methods can be viewed from a broader log linear model with latent variables perspective and how under that framework classification error can be better specified through greater model flexibility.

25 Spatial Variation and Risk Factors for Disease ●

Section on Statistics in Epidemiology, Section on Statistics and the Environment, Section on Risk Analysis, Biometrics Section

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Risky Behavior: A Multivariate Statistical Analysis of the United States Based on Health Risk Factors

✱Ashley M. Swandby, SUMSRI (Miami University REU), 3891 Read Mountain Road, Cloverdale, VA 24077, *aswandby@gmail.com*; Christina McIntosh, SUMSRI (Miami University REU); Alicia Smith, SUMSRI (Miami University REU)

Key Words: Multivariate Statistics, Undergraduate REU, health risk factors, Principal Component Analysis, Factor Analysis, Discriminant Analysis

In this paper, we study a number of variables associated with health risk factors in the United States using the 2006 Centers for Disease Control and Prevention's Behavioral Risk Factor Surveillance System survey data. We use Principal Component Analysis, Factor Analysis, and Discriminant Analysis in order to analyze the multivariate data. Furthermore, we provide a ranking of relative health for some of the states based on the analysis.

Is It Race/Ethnicity, or Is It Socioeconomic Status? Disparities in the Prevalence of Diabetes in the Boston Area Community Health (BACH) Survey

✱Carol L. Link, New England Research Institutes, 9 Galen Street, Watertown, MA 02472, *clink@neriscience.com*; John B. McKinlay, New England Research Institutes

Key Words: Diabetes, Socioeconomic status, epidemiology

Minorities (blacks and Hispanics) are reported to be at increased risk of diabetes compared to whites. In a community-based random sample of Boston residents aged 30 to 79, we find that blacks (odds ratio 2.04, p=.0001) and Hispanics (odds ratio 2.35, p<.0001) have more than twice the odds of diabetes compared to whites after adjusting for gender and age. However, once socioeconomic status (measured as a combination of education and income), body mass index (BMI), alcohol consumption, and trouble paying for basics are also added to the model, these odds ratios drop to 1.32 (p=.1275) and 1.42 (p=.0902), respectively. These results suggest that much of the race/ethnic variation in the prevalence of diabetes in the United States may be due to modifiable socioeconomic factors. These findings have profound implications for social policy, public health interventions, and clinical practice.

Estimation for Age-Period-Cohort Models: With Application to Mesothelioma Data in Alberta

✱Bei Jiang, University of Alberta, 632 Central Academic Building, University of Alberta, Edmonton, AB T6G 2G1 Canada, *bjiang@math.ualberta.ca*; K.C. Carriere, University of Alberta

Key Words: Age-Period-Cohort Model, Log-linear model, Estimable functions, Regression splines, Smoothing splines, Generalized Additive Models

Age-Period-Cohort (APC) models have been widely used by epidemiologist to analyze the trends of disease incidence and mortality. The linear dependency: "Cohort=Period-Age" results in the well-known "nonidentifiability" problem. We review the "solutions" built on estimable functions (Holford, 1983). We show how the restricted regression splines can be used to extend the classical framework of factorial APC models (Heuer 1997). Further, we consider the use of Generalized Additive Models (GAM) for the estimation and present how the "non-identifiability" problem can be dealt with accordingly. Finally, these methods are illustrated to the Mesothelioma data in Alberta.

A Model for Spatio-Temporally Clustered Disease Rates

✱Ronald Gangnon, University of Wisconsin, 603 WARF Office Building, 610 Walnut St, Madison, WI 53726, *ronald@biostat.wisc.edu*

Key Words: breast cancer, spatial modeling, temporal modeling, cluster detection, Bayes factor

We describe an extension of a model for spatial clustering proposed by Gangnon and Clayton (2001) to spatio-temporal data. As in the purely spatial model, a large set of circular regions of varying radii centered at observed locations are considered as potential clusters (e.g., subregions with a different pattern of risk than the remainder of the study region). Within the spatio-temporal model, no specific parametric form is imposed on the temporal pattern of risk within each cluster. In addition to the clusters, the proposed model incorporates spatial and spatio-temporal heterogeneity effects and can readily accommodate regional covariates. Inference is performed in a Bayesian framework using MCMC. We illustrate the approach with an application of the model to data on female breast cancer mortality in Japan.

Geocoding Accuracy: Effects of Geographic Features and Spatial Autocorrelation

✱Jie Li, The University of Iowa, 213 Schaeffer Hall, The University of Iowa, iowa city, IA 52242-1409, *jie-li-1@uiowa.edu*; Dale Zimmerman, The University of Iowa

Key Words: E911, geocoding error, orhotphoto-based geocodes, regression, TIGER

We investigated two main research questions pertaining to geocoding accuracy of residential addresses using a dataset of 9298 geocoded addresses from Carroll County, Iowa. Three distinctly different geocodes were obtained: TIGER, E911, and orthophoto-based geocodes, with the latter regarded as the truth. Positional errors associated with TIGER and E911 geocoding were thus determined. Geographic characteristics associated with each address were also determined, some of which were found to have statistically significant effects on the magnitudes of geocoding errors. Furthermore, treating the positional error of each standard geocode as a spatially indexed attribute variable, we found highly significant evidence of spatial autocorrelation for this variable. The implications of these findings for statistical modeling of geocoded data are discussed.

A Comprehensive Multilevel Model of Infant Mortality: A Case Study in Shelby County, Tennessee

✱Sudip Chatterjee, The University of Memphis, 619 Patterson Street, Apartment #7, Memphis, TN 38111, *schattrj@memphis.edu*

Key Words: Social Model, Multilevel Model, GLIMMIX, Medical Model, Infant Mortality, Spatial

For the last five years, Shelby County has recorded more than 12 infant deaths per 1,000 live births, as compared to eight per 1,000 live births in the state of Tennessee during the same period, which suggest that there is a systematic problem in Shelby County. Several models have been developed to understand the determinants of infant mortality. These models are classified as social models and medical models. This study integrates a social model and medical model of infant mortality to ascertain the most important predictors of infant mortality and to determine how these predictors vary spatially. A multilevel modeling approach was adopted using Glimmix (SAS Macro). My model estimates that there is an effect of exposure to domestic violence (in the neighborhood) on the proximate determinant of low gestational period, which then leads to high infant mortality in Shelby County, TN.

A Generalized Linear Models Approach to Spatial Scan Statistics for Covariate Adjustment

✱Inkyung Jung, University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, Mail Coe 7933, San Antonio, TX 78229, *inkyung.jung@gmail.com*

Key Words: cluster detection, confounding factor, geographical disease surveillance, GLM

The spatial scan statistic proposed by Kulldorff (1997) is one of the most widely used methods for detecting spatial clusters and evaluating their statistical significance. However, it is not fully capable of adjusting for all types of confounding covariates. In this article, a generalized linear models (GLM) approach to construct spatial scan statistics, which is readily in a form for covariate adjustment, is proposed. Using GLM, spatial scan statistics for different probability models can be formulated in a single framework. The test statistic is based on the log-likelihood ratio test and evaluated using Monte Carlo hypothesis testing. The proposed method is illustrated using Texas female breast cancer data concerning late versus early stage cancer cases with covariates of race/ethnicity and age group.

26 Applications of Survival Time Models ●

Section on Statistics in Epidemiology, Section on Quality and Productivity, Biometrics Section

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Joint Modeling of Survival and Multinomial Data with Applications to Prostate Cancer Stage-Grade Specific Incidence

✱Chen Hu, The University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109-2029, *chenhu@umich.edu*; Alexander Tsodikov, The University of Michigan

Key Words: Cancer Incidence, Tumor Progression, Compartmental Model, Screening

The usage of prostate-specific antigen (PSA) provides early detection of prostate cancer and exerts a remarkable effect on cancer incidence. Stage-grade specific incidence represents a random vector of age at diagnosis, and cancer stage and grade, with a joint survival and multinomial distribution. A series of survival models with time-dependent covariates are considered. Our modeling focuses on discrimination between two hypotheses regarding the mechanism of grade progression—selection and adaptation. We study model identifiability and develop model-based discrimination methodology for the problem. Random screening schedules are incorporated into the model to represent the effect of PSA tests. The models are applied to data from the Surveillance, Epidemiology, and End Results (SEER) program.

Time to Recurrence of Shingles (Herpes Zoster) Infection

✱Peter C. Wollan, Olmsted Medical Center, Research Department, 210 9th St SE, Rochester, MN 55904, *pwollan@olmmed.org*; Patricia Saddier, Merck & Co., Inc.; Barbara P. Yawn, Olmsted Medical Center

Key Words: multiple-event survival analysis, Cox regression

Yawn et al (Mayo Clinic Proceedings, 2007) reported a population-based study of herpes zoster (HZ) incidence for the years 1996-2001, funded by Merck and Co., Inc. As a follow-up study, the medical records of the 1669 HZ index cases were re-examined to identify HZ recurrences. New episodes after the index episode were found in 95 subjects, with 8 subjects having more than one. Overall, the 8-year recurrence rate was 6.6% (Kaplan-Meier estimate). Females and those with impaired immune status had significantly higher risk. Older subjects had a slightly higher recurrence rate, but the difference was not statistically significant. In addition, 93 events were identified prior to the HZ index episode. We discuss alternative approaches to analyzing these observations, including the conditional multiple-event model, basing time on subject's age, and basing time on first event.

Assessment Predictive Power of Prognostic Models for Melanoma Survival Data

✱Shouluan Ding, The University of Alabama at Birmingham, 1530 3rd Anenue South, WTI 153, Birmingham, AL 35294, *sding@uab.edu*; Seng-jaw Soong, The University of Alabama at Birmingham

Key Words: Predictive Power of Model, Concordance Correlation coefficient, Melanoma survival Data, Extended generalized Gamma Distribution, Parametric Models, Cox model

We have developed the predictive Cox model and Extended Generalized Gamma (EGG) model in melanoma research. For model validation, which is the most overlooked, we used AJCC melanoma collaborative database divided into a training dataset with 14,760 patients from 10 institutions for Cox and EGG developing of models by thickness 7 subgroups and a test dataset with 10,974 patients from Sydney Melanoma Unit (SUM) for validating by the same groups. Concordance correlation coefficient (CCC) was used to assess the predictive power of Cox and EGG by comparing predicted survival rates with Kaplan-Meier observed survival rates at 5- and 10-year time point for individual patients with combination of significant clinical or pathologic prognostic factors within thickness groups for SUM dataset. Overall CCC of 5- and 10-year survival rates were excellent consistency, at 0.87 and 0.89, respectively.

Estimating Prostate Cancer Survival Under a Misattribution of the Cause of Death

✱Jinkyung Ha, The University of Michigan, 2925 N. Knightsbridge Cir., Ann Arbor, MI 48105, *jinha@umich.edu*; Alexander Tsodikov, The University of Michigan

Key Words: misattribution, competing risk, cancer mortality

The observed trend of prostate cancer mortality over the period of increased utilization of screening follows that of cancer incidence. However, changes in the incidence due to screening are expected to lag those of mortality as patients with prostate cancer typically have good prognosis. One of possible explanation for the observed trend could be misattribution of the underlying cause of death in prostate cancer patients. In this work, we propose a nonparametric maximum likelihood estimation (NPMLE) approach for the post-treatment survival under misattribution, and derive isotonic MLE of the underlying survival function. We then explore small-sample and asymptotic properties of the procedure. The method is applied to data from the Surveillance, Epidemiology and End Results (SEER) program. We compare our method with other solutions to similar problems proposed earlier.

A Nonparametric Approach for Comparing Survival Distributions with Propensity Score Adjustment

✱Nandita Mitra, University of Pennsylvania, Department of Biostatistics and Epidemiology, 929 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, *nanditam@mail.med.upenn.edu*; Katrina Armstrong, University of Pennsylvania; Glenn Heller, Memorial Sloan-Kettering Cancer Center; Venkatraman E. Seshan, Columbia University

Key Words: propensity score, survival data, nonparametric test statistic, Medicare

We propose a nonparametric test for the equality of the conditional hazard function between comparison groups [Heller and Venkatraman, 2004] that allows for the adjustment of concomitant variables via the estimated propensity score. This conditional logrank test does not require the assumption of proportional hazards, a parametric specification of the relative risk function, or randomization of group assignment. The methodology is demonstrated in a large observational study of 76,286 Medicare enrolled prostate cancer patients that compares the survival benefit of men who received prostatectomy versus those who did not. Numerous covariates such as race, age, socio-economic status, and comorbidities were adjusted for via the propensity score.

Evaluation of Prediction in Models of Composite Endpoints and Their Components

✱Robert Glynn, Brigham & Women's Hospital, Division of Preventive Medicine, 900 Commonwealth Avenue, Boston, MA 02215, *rglynn@rics.bwh.harvard.edu*; Bernard Rosner, Channing Laboratory

Key Words: prediction, C-statistic, competing risks, survival

Composite endpoints are commonly used in both observational and randomized studies for a broad perspective on the impact of exposures of interest. In the setting of time to occurrence of outcomes, one can use methods of competing risks survival analysis to evaluate possible heterogeneity of associations of exposures with different components of the composite endpoint. We present two alternative approaches to quantify discrimination in such competing risks models. One relative approach considers likelihood ratios from nested models that assume common vs. distinct associations of exposures on different components of the endpoint. A second approach generalizes the C-statistic proposed by Harrell to the setting of composite endpoints. Approaches are illustrated with data on a composite endpoint including first occurrence of arterial or venous thrombosis in a prospective cohort study.

27 Topics in Statistical Testing

IMS, Section on Nonparametric Statistics
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Hazard-Based, Data-Driven, Goodness-of-Fit Tests

✱Ma. Zenia Agustin, Southern Illinois University Edwardsville, Department of Mathematics and Statistics, Box 1653, Edwardsville, IL 62026, *zagusti@siue.edu*

Key Words: Neyman's test, score test

In this talk, we will explore the goodness-of-fit problem involving a composite hypothesis testing scenario. In particular, a family of hazard-based smooth goodness-of-fit tests will be considered. This family of tests is applicable to survival and reliability models, as well as other models formulated in terms of counting processes. Of particular interest is the determination of an appropriate smoothing parameter for this family of tests. The results of a Monte Carlo simulation will be presented to illustrate the potential of these tests as powerful omnibus and directional tests.

EDF Goodness-of-Fit Tests for Testing the Distributional Assumptions in ANOVA

✱Dhanuja Kasturiratna, Northern Kentucky University, Department of Mathematics, Northern Kentucky University, Highland Heights, KY 41099, *kasturirad1@nku.edu*; Truc T. Nguyen, Bowling Green State University; Arjun K. Gupta, Bowling Green State University

Key Words: ANOVA, Transformations, Goodness-of-fit tests

In the one-way classification with k treatments, we assume that the data are observed according to the linear additive model, where the error random variables are independently and identically distributed normally with mean zero and constant variance. Then to test whether set of observed data are coming from the above regression model, we need to construct a test for testing the hypothesis that the observations of the treatment groups are normally distributed with the same variance. In this paper, the location and scale parameters in the distributional assumption of the regression model are eliminated by means of suitable transformations on the original observations. The transformed random variables are then shown to possess Student's t distributions with known parameters under the null hypothesis. The goodness-of-fit tests can then be performed on the transformed variables.

Goodness-of-Fit Tests for Multivariate Normality

✱Ming Zhou, Iowa State University, Department of Statistics & Statistical Laboratory, Snedecor Hall, Ames, IA 50011-1210, *zhouming@iastate.edu*

Key Words: multivariate normal, goodness-of-fit, power, asymptotic optimality, location scale invariant

Multivariate normal distribution is one of the most widely used multivariate models in statistics. In this talk we will discuss some new nonparametric goodness of fit tests for assessing multivariate normality. Both large sample optimality properties of the proposed tests and Monte Carlo simulation results for comparative power performance will be presented. This is based on the joint work with Dr. Yongzhao Shao.

Asymptotic Expansion of the Null Distributions of Test Statistics for Profile Analysis Under General Conditions

✱Solomon W. Harrar, University of Montana, Department of Mathematical Sciences, 32 Campus Drive, Missoula, MT 59812, *harrar@mso.umt.edu*; Jin Xu, East China Normal University

Key Words: Asymptotic Expnasion, Bartlett's Correction, Cumulants, Multivariate Tests, Nonnormality, Profile Analysis

In this talk, we present asymptotic expansion results on the null distributions of some test statistics for k-sample profile analysis under general conditions. Furthermore, new statistics are proposed for testing the hypothesis of flatness in view of the fact that this test is conducted following acceptance of the parallelism hypothesis. Numerical accuracy of the results is ivestigated via a simulation study. A real data example illustrates the application.

Conditioning on the Tail: A Rationale for Using Traditional Two-Tailed Tests and Confidence Intervals

Paul W. Vos, East Carolina University; ✱Suzanne Hudson, East Carolina University, Dept. of Biostatistics, Allied Health, East Carolina University, Greenville, NC 27858, *hudsons@ecu.edu*

Key Words: two-tailed tests, confidence intervals, conditioning

The P-value for a traditional two-tailed hypothesis test is twice the smaller tail probability. We consider only exact tests, i.e., those that compute the P-value using the true distribution, not some approximation to it. We also consider only one-parameter distributions. Some statisticians consider a two-tailed test to be appropriate for symmetric distributions, but not for skewed distributions. We argue that one should condition on the tail where the data falls. The P-value computed from such a conditional test is the same as the two-tailed P-value. It is shown how a test that conditions on the tail can be constructed. Problems with alternative exact tests for discrete distributions and the associated confidence intervals are discussed.

Moderate Deviations for Two Sample T-Statistics

✱Hongyuan Cao, The University of North Carolina at Chapel Hill, 881 Martin Luther King Jr. Blvd Apt. 18H, Chapel Hill, NC 27514, *hycao@email.unc.edu*

Key Words: two sample t-statistics, asymptotic distribution, moderate deviation

Let $X_1, \cdots, X_{n_{1}}$ be a random sample from a population with mean μ_1 and variance σ_1^2, and $Y_1, \cdots, Y_{n_{2}}$ be a random sample from another population with mean μ_2 and variance σ_2^2. The sampling process is independent of each other. Consider the two sample t-statistic $ T={{\bar X-\bar Y-(\mu_1-\mu_2)} \over \sqrt{s_1^2/n_1+s_2^2/n_2}}$. This paper shows that $ \ln P(T \geq x) \sim -x^2/2 \ $ for any $x:=x(n_1,n_2)$ satisfying $x \to \infty$, $x=o(n_1+n_2)^{1/2}$ as $n_1, n_2 \to \infty$ provided $ 0 < c_1 \leq n_1/n_2 \leq c_2 < \infty.$ If, in addition, $E|X_1|^3 < \infty$, $E|Y_1|^3 < \infty$, then $\frac{P(T \geq x)}{1-\Phi(x)} \to 1 $ holds uniformly in $x \in(0,o((n_1+n_2)^{1/6}))$.}

Concomitants of Order Statistics for Bivariate Pseudo Inverse Rayleigh Distribution

✱Muhammad Q. Shahbaz, COMSATS Institute of Information Technology, Department of Mathematics, Lahore, International 54000 Pakistan, *qshahbaz@gmail.com*; Munir Ahmad, NCBA&E

Key Words: Order Statistics, Concomitants, Rayleigh Distribution, Pseudo Inverse

Order statistics has been widely used in distribution theories. Concomitants have a vital role in study of arranged sample from a bivariate population. In this paper the distributional properties of concomitants of order statistics from bivariate pseudo inverse Rayleigh distribution have been discussed. The single and product moments of the concomitants have been obtained.

28 Monte Carlo Methods

Section on Statistical Computing, IMS
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Automated Markov Chain Monte Carlo Based on the Ratio of Uniforms Transformation

✱Chris Groendyke, The Pennsylvania State University, 3101 Shellers Bend, State College, PA 16801, *cxg928@psu.edu*

Key Words: Markov chain Monte Carlo, Ratio of uniforms, Auxiliary variable methods, Linear Gaussian process models

Markov chain Monte Carlo (MCMC) algorithms are commonly used for sampling from posterior distributions in Bayesian statistics. These algorithms often require a lot of tuning, which can be time consuming and difficult for non-experts. I use the ratio of uniforms (ROU) transformation due to Kinderman and Monahan to develop several MCMC algorithms which require little or no tuning, yet still work well in practice. I demonstrate the application of these algorithms to sampling from the posterior distribution of a linear Gaussian process model, a popular model for spatial statistics. I conclude with a comparison of the performance of several automated algorithms in the context of real data sets.

The Monte Carlo Fisher Scoring Method

✱Yong Wang, The University of Auckland, Department of Statistics, The University of Auckla, Private Bag 92019, Auckland, 1142 New Zealand, *yongwang@stat.auckland.ac.nz*

Key Words: Fisher scoring, iteratively reweighted least squares, EM algrithm, mixture model, Monte Carlo integration, randomized algorithm

We present a simulation-based Fisher scoring method for maximum likelihood estimation of parameters, with or without constraints. The method is particularly useful in the multivariate situation when the expected Fisher information matrix does not have a closed-form expression. Specifically, random samples are generated from the estimated models and used to provide sample estimates of the Fisher information matrices. The log-likelihood function is then iteratively approximated by quadratic functions that have the same gradients at the iterates but random Hessians. It is shown that each quadratic can be readily optimized by solving a linear regression problem. We apply the method to fit multivariate mixture models, and numerical studies show that it outperforms, often remarkably, the conventional EM algorithm.

On the Assessment of Monte Carlo Error in Statistical Experiments

✱Elizabeth Koehler, Vanderbilt University, 571 Preston Research Building, Department of Biostatistics, Nashville, TN 37232-6848, *elizabeth.koehler@vanderbilt.edu*; Sebastien Haneuse, Group Health Center for Health Studies; Elizabeth R. Brown, University of Washington

Key Words: Monte Carlo, Monte Carlo error, Replications, Simulation Study

Statistical experiments, more commonly referred to as Monte Carlo or simulation studies, are used to study the behavior of statistical methods under controlled situations. While recent computing and methodological advances have permitted increased efficiency in the simulation process, such experiments are limited by their finite nature and are subject to uncertainty. There has been, however, virtually no emphasis placed on reporting the uncertainty, or Monte Carlo error, associated with simulation results in the published literature, or on justifying the number of replications used. We present a series of simple, practical methods for estimating Monte Carlo error and the number of replications required to achieve a desired level of accuracy. The methods are demonstrated in example simulations. Our results suggest Monte Carlo error may be more substantial than traditionally thought.

Application of a New Multivariate Resampling Method To Improve Statistical Performance of Multiple Regression with Small Samples

✱Haiyan Bai, University of Central Florida, *hbai@mail.ucf.edu*; Wei Pan, University of Cincinnati

Key Words: multiple regression, resampling, standard error, parameter estimate

The issues of estimation accuracy and statistical power in multiple regression with small samples have long been a concern. To improve the statistical performance, the bootstrap has been actively used through resampling the residuals; however, if the residuals are not normally distributed before resampling, the bootstrap usually fails (Freedman, 1981). The present study utilizes a new multivariate resampling method, the Sample Smoothing Amplification Resampling Technique (S-SMART), to improve the estimation accuracy and statistical power in multiple regression with small samples. S-SMART is a distribution-free method and employs smoothing to create multivariate resamples based on given small sample. The S-SMART resamples possess the same statistical properties as does the small sample. The statistical performance of multiple regression will be improved through the S-SMART procedure.

Phylogenetic Consensus Tree Construction Using Stochastic Approximation Monte Carlo

✱Sooyoung Cheon, University of Virginia, Department of Public Health Sciences, West Complex, RM 3181, Charlottesville, VA 22908-0717, *sc8te@virginia.edu*

Key Words: Bayesian phylogeny, Local trap, Markov chain Monte Carlo, Stochastic approximation Monte Carlo

Monte Carlo methods have received much attention recently in the literature of phylogeny estimation. However, they often suffer from the local-trap problem in simulations. In this paper, we apply stochastic approximation Monte Carlo (SAMC) to the problem. SAMC effectively prevents the system from getting trapped in local energy minima. SAMC is compared with a variety of existing Bayesian and non-Bayesian methods for constructing consensus trees on simulated and real datasets. Numerical results favor to SAMC, which tends to produce more accurate estimates for the evolutionary parameters and the probability of each edge being included in the phylogenetic trees than do the existing methods.

Public Hurricane Loss Evaluation Models: Predicting Losses of Residential Structures in the State of Florida

Shahid Hamid, Florida International University; ✱B.M. Golam Kibria, Florida International University, Department of Statistics, Miami, FL 33199 USA, *kibriag@fiu.edu*; Sneh Gulati, Florida International University; Mark Powell, NOAA Hurricane Research Division; Bachir Annane, CIMAS, Rosenstiel School/University of Miami; Steve Cocke, Florida State University; Jean-Paul Pinelli, Florida Institute of Technology; Shu-Ching Chen, Florida International University

Key Words: Damage, Goodness-of-fit, Hurricane, Sensitivity & Uncertainty, Validation, Wind Model

As an environmental phenomenon, hurricanes cause significant property damage and loss of life in coastal areas almost every year. Although a number of commercial loss projection models have been developed to predict the property losses, only a handful of studies are available in the public domain to predict damage for hurricane prone areas. The State of Florida has developed an open, public model for the purpose of probabilistic assessment of risk to insured residential property associated with wind damage from hurricanes. The model comprises of atmospheric science, engineering & actuarial components. At every stage of the process, statistical procedures were used to model various parameters and validate the model. This paper presents a brief summary of the main components of the model (meteorology, vulnerability & actuarial) and then focuses on the statistical validation of the same.

29 Applications in Health Promotion

Section on Health Policy Statistics, Section on Statistics in Epidemiology
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Multilevel Analysis for Child Nutrition and Health: Pick a Better Snack in Iowa

✱Mack Shelley, Iowa State University, 323 Snedecor Hall, Ames, IA 50011-1210, *mshelley@iastate.edu*; Doris Montgomery, Iowa Department of Public Health; Xia Chen, Iowa Department of Public Health

Key Words: health, nutrition, physical activity, multi-level analysis, public policy, school-based

The Pick a Better Snack program provides school-based education about the benefits of nutrition and physical activity. In Iowa, this information is provided by the Iowa Nutrition Network's social marketing campaign to encourage fruit and vegetable choices for snacks, with support from public and nonprofit agencies. Multilevel models using data collected from multiple school sites in Iowa are estimated to predict self-reported student physical activity, health and nutrition knowledge, self-efficacy, and fruit and vegetable consumption. We examine the level-2 effects of differences in projects, teachers, urbanicity, program intensity, and socioeconomic status (measured by the percentage of students eligible for free and reduced lunch) and the level-1 effects of student differences based on grade level and demographics. Implications for public nutrition policy are discussed.

Simpson's Paradox in Real Life: The Center for Medicare Services Composite Quality Score

✱Carlos F. Alzola, Data Insights, 401 Glyndon St SE, Vienna, VA 22180, *calzola@cox.net*; Sam Ogunbo, Maryland Hospital Association; Nikolas Matthes, Maryland Hospital Association

Key Words: Simpson's paradox, CMS, Pay for performance, Scores, Medicare

The National Hospital Quality measures (NHQM) assess different dimensions of clinical conditions. Quality of care for AMI patients is assessed with nine different indicators along a continuum of care. The indicators, taken individually, do not provide an overall picture of the care provided by a given facility. CMS's composite score looks at each patient and each dimension of care as an opportunity to provide eligible care. All these opportunities are aggregated over all patients at a facility to form a denominator. The opportunities where patients received eligible care are also aggregated to form a numerator. The ratio of these two quantities form the Composite Quality Score. The methodology is susceptible to instances of Simpson's Paradox. We explore how often hospitals that perform better than another in each individual measure are ranked lower when compared on their CQS scores.

Analysis of Health and Safety Data: Near Misses and Injury-Illness Rates

✱Rama Sastry, U.S. Department of Energy, HS-32, Washington, DC , *Rama.sastry@hq.doe.gov*

Key Words: Health, Safety, Near Misses, Injury Rates

The Department of Energy (DOE) is responsible for protecting health and safety of 130,000 workers at several sites such as Hanford, Savannah River Site, and the National Laboratories at Oak Ridge, Los Alamos, etc. The health and safety data reported by the DOE contractors is maintained by two large data bases, Occurrence Reporting Processing System (ORPS) and Computerized Accident/incident Reporting System (CAIRS). The injury-illness data are consistent with Occupational Safety and Health Administration (OSHA) reporting regulations. The objective of this paper is to present some results of the application of statistical techniques to combine the elements of the two separate data bases, for example the Near Misses data from ORPS and injury rates from CAIRS.

Evaluating the Sizes of Differences Between Group Rates in Settings of Different Overall Prevalence

✱James P. Scanlan, James P. Scanlan, Attorney at Law, 1529 Wisconsin Ave., NW, Suite 300, Washington, DC 20007, *jps@jpscanlan.com*

Key Words: health disparities, health inequalities, measurement, binary measures

As prevalence of an outcome changes, binary measures of differences between rates of experiencing or avoiding it tend to change as well. As an outcome like mortality declines in prevalence, relative differences between rates of experiencing it tend to increase while relative differences between rates of avoiding it tend to decline. Absolute differences between rates and odds ratios also tend to change, though in a more complicated manner. As the outcome changes from being nearly universal to being rare, absolute differences tend to increase for a time and then decline, while odds ratios decrease for a time and then increase. This session would explain these tendencies and the obstacles they create for appraising the size of group differences in settings of different prevalences, and explore two approaches to measuring group differences that avoid those obstacles.

Analysis of National Survey Data on Public Concerns About Avian Influenza for Preparedness Planning

✱William E. Pollard, Centers for Disease Control and Prevention, 1600 Clifton Road, NE, Mailstop E-69, Atlanta, GA 30329, *bdp4@cdc.gov*

Key Words: health surveys, public health communication, avian influenza

Communicating with the public about preparedness for avian influenza is an important function of the Centers for Disease Control and Prevention (CDC), and this requires an understanding of the audience. This paper describes analysis of 2006 and 2007 proprietary national survey data from the Porter Novelli ConsumerStyles survey (12,000 respondents per year) with lifestyle and media habit content, along with items about concern with avian flu and attention to media coverage, and from the Porter Novelli HealthStyles survey, (4,000 respondents per year who have also completed the consumer survey) with health attitude and behavior content. The linked data sets provide extensive information for understanding factors relating to concern and attention. Chi-squared automatic interaction detection (CHAID) techniques are used to identify different audience segments for communication planning.

Patient Teenagers: Virginity Pledges as a Marker for Lower Sexual Activity

✱Janet Rosenbaum, University of Illinois at Chicago, 3530 N Broadway St #3, Chicago, IL 60657, *janet@post.harvard.edu*

Key Words: public health, abstinence, adolescent health, matched sampling, propensity scores, sexual behavior

The US spends $200 million annually on abstinence programs, including virginity pledges. Using data from Add Health waves 1–3, adolescents reporting a wave 2 virginity pledge (n=289) were matched with nonpledgers (n=645) using exact and nearest-neighbor matching on wave 1 factors including pre-pledge religiosity, attitudes towards sex and birth control, and family context. Wave 3 outcomes were compared. Five years post-pledge, 84% of pledgers reported having never taken a pledge; pledgers and matched nonpledgers did not differ in premarital sex and sexually transmitted diseases. Pledgers had 0.1 fewer past year partners, but the same number of lifetime sexual partners and same age of first sex. Pledgers were 10 percentage points less likely than matched nonpledgers to use condoms in the last year, and less likely to use birth control in the past year and at last sex.

Estimation of Net Present Value of Total Health Care Costs

✱Lin Liu, Eli Lilly and Company, Indianapolis, IN 46285, *liull@lilly.com*; Joseph C. Gardiner, Michigan State University; Zhehui Luo, RTI International; Charles Given, Michigan State University

Key Words: Markov Model, Transition Intensity, Health States

In follow-up studies of health care costs, patients may transit between health states and incur costs at different rates while sojourning in these states. We propose a method for estimating total health care costs that combines a Markov model for transitions between states and a longitudinal regression model for costs. In our model costs emanate from two streams, during sojourn in health states and in transition between states. We estimate net present value (NPV) for costs incurred over a fixed period and derive the asymptotic properties of the estimator. The method is applied to a data set on cancer patients to obtain NPV estimates of total health care cost by cancer stage. Resampling methods provide standard errors. Simulation studies demonstrate that the estimator performs well and can be applied in moderately large data sets with health care costs assessed over time.

30 Bayesian Models and Methods for Spatial and Temporal Data ●▲

Section on Bayesian Statistical Science, Section on Statistical Computing

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Change of Spatiotemporal Scale in Dynamic Models

✱Yongku Kim, Statistical and Applied Mathematical Sciences Institute, 19 T.W. Alexander Drive, Research Triangle Park, NC 27709, *kim@samsi.info*; Mark Berliner, The Ohio State University

Key Words: Change of support, Bayesian inference, spatiotemporal model, time-varying parameter

Spatiotemporal processes show complicated and different patterns across different space-time scales. Both scientific understanding and observed data vary in form and content across scale. Such information sources can be combined through Bayesian hierarchical framework. This approach restricts

a few essential scale. However, it is common in the trade-off view between simple modeling and analysis strategy with complicate modeling. We extended Wikle and Berliner (2005)'s hierarchical COS strategy by adding temporal modeling in their style and allowing discretized time-varying parameters. We apply a Bayesian inference based on combining information across spatiotemporal scale to some climate temperature data which are point referenced data and areal unit data. The inference focuses on the temperature process on specific prediction grid scale and maybe different time scale.

A Stochastic Neighborhood Conditional Auto-Regressive Model for Spatial Data

✱Gentry White, North Carolina State University, 201C Patterson Hall, Campus Box 8203 NCSU, Raleigh, NC 27695-8203, *white@stat.ncsu.edu*; Sujit Ghosh, North Carolina State University

Key Words: Spatial data analysis, conditional autoregressive models, geostatistical models

A spatial process over a lattice or a set of irregular regions can be modeled as a conditionally autoregressive (CAR) model. The neighborhoods within a CAR model are formed using the physical features of the regions. We propose a stochastic neighborhood within a CAR model, (SNCAR). This neighborhood is defined as an extension of a recently proposed model that can approximate other geostatistical models. The parameters determining the neighborhood measure are assigned priors and estimated within a hierarchical model. These parameters define a prior for spatial effects. The resulting model is flexible in accurately estimating covariance structures for a variety of data. Specific examples are illustrated using data generated from some common spatial covariance functions as well as real data concerning radioactive contamination of the soil in Switzerland after the Chernobyl accident.

Combining Complex Climate Models with Massive Observational Data for Predicting Climate Change

✱K. Sham Bhat, The Pennsylvania State University, 326 Thomas Building, Department of Statistics, University Park, PA 16802, *kgb130@psu.edu*; Murali Haran, The Pennsylvania State University; Klaus Keller, The Pennsylvania State University; Joshua Dorin, The Pennsylvania State University

Key Words: spatiotemporal data, Bayesian modeling, climate change, Gaussian process, kernel convolution

We study the risk of a collapse of the meridional overturning circulation (MOC), part of the global ocean circulation `conveyor belt,' a key factor determining global climate patterns. Assessing this risk involves combining massive computer climate model output with space-time intensive physical observations of the ocean system. We use a Bayesian approach to connect the two sets of data by approximating the climate model output using a Gaussian process based emulator and using Markov Chain Monte Carlo (MCMC) methods to obtain a posterior distribution for important parameters associated with the probability of an MOC collapse. We use a kernel convolutions approach (Higdon 1998) to make our approach computationally tractable.

Zero-Inflated Bayesian Spatial Models with Repeated Measurements

✱Jing Zhang, University of Missouri-Columbia, 510 High Street Apt.316, Columbia, MO 65201, *jztw7@mizzou.edu*; Chong Z. He, University of Missouri-Columbia

Key Words: MCMC, Spatial, Zero-inflated, Repeated measurements, Bayes factor

When there are so many zeros in the data that it does not readily fit any standard distributions, such as normal, binomial, Poisson or log-normal, the data set is refereed to as "zero-inflated." However, most of the literatures focused on the zero-inflated count data, while there is limited literature analyzing the zero-inflated continuous spatial data. This talk considers a zero-inflated Bayesian spatial model with repeated measurements which deals with the continuous data. The inference, including simulating from the posterior distributions, predicting on new locations as well as hypothesis testing on the model parameters, is implemented using Markov Chain Monte Carlo (MCMC) techniques. The methodology is also applied to the herbaceous data in the Missouri Ozark Forest Ecosystem Project. The analysis of multivariate (zero-inflated) spatial data is also of interest.

Gaussian Multiscale Spatio-Temporal Models

✱Marco A.R. Ferreira, University of Missouri-Columbia, 146 Middlebush Hall, Columbia, MO 65211-6100, *ferreiram@missouri.edu*; Scott Holan, University of Missouri; Adelmo Bertolde, Federal University of Espirito Santo

Key Words: space-time models, dynamic models, Bayesian inference, multiscale analysis

We develop a new class of multiscale spatio-temporal models for Gaussian areal data. Our framework decomposes the spatio-temporal observations and underlying process into several scales of resolution, and evolves the multiscale coefficients through time with structural state-space equations. The multiscale decomposition considered here is able to accommodate irregular grids and heteroscedastic errors. Our multiscale spatio-temporal framework has several salient attributes. First, the multiscale decomposition leads to an extremely efficient divide-and-conquer estimation algorithm. Second, the multiscale coefficients have an interpretation of their own. Finally, deterministic relationships between different resolution levels are automatically respected. We illustrate the use of our multiscale framework with an analysis of a spatio-temporal dataset on agriculture production in Brazil.

Bayesian Estimation of GARCH Models

✱Daniel Wan Hussin, University Bamberg, Feldkirchenstr. 21, Bamberg, 96052 Germany, *daniel.wan-hussin@uni-bamberg.de*

Key Words: Bayes, GARCH, Model Averaging, MCMC

Autoregressive conditionally heteroscedastic (ARCH) models introduced by Engle in 1982 and the generalization by Bollerslev are commonly used in financial statistics to model volatility. Since their introduction many generalizations of the GARCH model have been developed. The presentation shows how Markov Chain Monte Carlo methods can be used to estimate the parameters of univariate GARCH models and many of its generalizations (EGARCH, TGARCH, Absolut Value GARCH, etc.) from the Bayesian point of view. Different residual distributions (normal, t-distribution, GED) are considered. Also the Bayesian estimation of the parameters of the model introduced by Hentschel is shown which nests all the above-mentioned generalizations. Results of the different simulations are presented and compared with the results of maximum likelihood estimation.

A Bayesian Model for Multivariate Functional Data

✱Ori Rosen, University of Texas at El Paso, Bell Hall 221, Dept. of Mathematical Sciences, El Paso, TX 79968, *ori@math.utep.edu*; Wesley Thompson, University of Pittsburgh

Key Words: Functional data, Mixed-effects, Cubic splines

We propose a method for analyzing multivariate functional data with unequally spaced observation times that may differ among subjects. Fitting multivariate observations simultaneously rather than fitting each variable separately may be advantageous if the error terms corresponding to each variable are correlated. Our method is formulated as a Bayesian mixed-effects model in which the fixed part corresponds to the mean functions, and the random part corresponds to individual deviations from these mean functions. Covariates can be incorporated into both the fixed and the random effects. The methodology is studied by simulation and illustrated with real data.

31 Statistical Methods for Clinical Trials I

Biometrics Section, Biopharmaceutical Section
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Approaches to Handling Data when Trial Conduct Deviates from the Prespecified Simon's Two-Stage Design

✱Yujun Wu, sanofi-aventis, 312 Hana Road, Edison, NJ 08817, *yujun.wu@sanofi-aventis.com*; Weichung Shih, University of Medicine and Dentistry of New Jersey

Key Words: Phase IIA cancer trials, Simon's two-stage designs, conditional probability, p-value

Simon's 'optimal' and 'minimax' two-stage designs are common methods for conducting phase IIA studies investigating new cancer therapies. However, these designs are rather rigid in their settings because of the pre-specified rejection rules and fixed sample sizes at each stage. In practice, we often encounter the problem that a study is unable to adhere to the event number and sample sizes of the original two-stage design. In this paper, we consider some approaches in handling situations where deviations or interruptions from the original Simon's two-stage design occur because recruitment of patients is slower than expected. We consider four scenarios and use conditional probabilities to address the issues commonly inquired by the scientific review board. We also discuss how to report p-values in these situations.

Approximately Optimal Continuous Stopping Boundaries in a One-Sided Standard Sequential Test

✱Lan Nygren, Rider University, 2083 Lawrenceville Rd, Lawrenceville, NJ 08648, *lnygren@rider.edu*

Key Words: Optimal boundaries, Average stopping time, Brownian motion, Boundary crossing, Stopping rule

In clinical trials, continuous monitoring is required when a large but nonsignificant test statistic is observed at an interim analysis. We consider a one-parameter family of continuous stopping boundaries that are optimal in the sense that they minimize the average stopping time. These boundaries also satisfy the required overall Type I error rate and power. The problem is formulated as a constrained optimization problem and is solved numerically using the differential evolution algorithm. The main results are the ready-to-use approximately optimal continuous stopping boundaries for a wide range of values of power and the most commonly used overall levels of significance.

Sample Size Calculations in Logistic Regression: Re-Revisited

✱Marepalli B. Rao, University of Cincinnati, Center for Genome Information, 3223 Eden Avenue, Cincinnati, OH 45267, *marepalli.rao@uc.edu*; Mohammed K. Alam, Forest Research Institute

Key Words: Binary covariate, Binary response variable, Wald Statistic, Nuisance parameter, Maximum Likelihood, Sample size

The research in this presentation is inspired by the paper - Sample Size Determination for Logistic Regression Revisited - penned by Demidenko published in *Statistics in Medicine*, Vol. 26, 2007, pp.3385-3397. The basic set-up is the Logistic Regression with a single covariate. The null hypothesis of interest is that the covariate has no impact on the binary response variable. The issue at heart is what should be the sample size for given level and power. The sample size formula will depend on the choice of a test statistic. A popular choice is Wald's statistic. Even for this statistic there are a number of choices what constitutes the standard error of the MLE of the regression parameter. The standard error can be evaluated at the null hypothesis or at the MLEs of the parameters. The formulas will be different. In this talk, we will explore systematically the differences that ensue.

Comparison of Power Between Randomized Discontinuation Design (RDD) and Upfront Randomization Design (RD) on Progression-Free Survival

✱Pingfu Fu, Case Western Reserve University, 10900 Euclid Avenue, Epidemiology and Biostatistics, Cleveland, OH 44106, *pxf16@case.edu*; Afshin Dowlati, Case Western Reserve University; Mark Schluchter, Case Western Reserve University

Key Words: study design, power comparison, clinical trial, progression free survival

For target-based cytostatic agents in oncologic drug development, single-arm phase II trials may not be appropriate. The RDD (an "enrichment" approach) has been proposed as an alternative to the RD for screening targeted agents. With the RDD, all enrolled patients are treated with an experimental agent for a fixed course of therapy, and those with stable disease or with clinical benefit are then randomized to either continue or discontinue treatment with the agent. Assuming exponential tumor growth where the effect of treatment follows the growth rate cutoff (GRC) model or the sensitive fraction (SF) model of Freidlin and Simon, simulations demonstrate that when time-to-event is the endpoint the RDD is more efficient than the RD for GRC model and SF model when the treatment effect is small; RD, however, has more power than RDD under SF model when the treatment effect is moderate-large.

Semiparametric Regression Analysis for Short-Term and Long-Term Covariate Effect with Survival Data

✱Rajeshwari Sundaram, National Institutes of Health, 6100 Executive Blvd., 7B03Q, Rockville, MD 20852, *sundaramr2@mail.nih.gov*; Song Yang, National Heart, Lung, and Blood Institute

Key Words: Lifetime, Censoring, Likelihood, Changepoint

We consider a new semiparametric regression model for survival data by separating the short term and long term effect on hazard ratio of the covariates when appropriate. In the new model, for each covariate the regression parameters reflect its short term and long term effect respectively. In addition to containing the classical proportional hazards model, the new model also contains the proportional odds model. Furthermore, it is flexible enough to allow the cases where hazard ratios between different values of a covariate could possibly cross. This work extends the two sample case studied by Yang and Prentice (2005). We use a pseudo-maximum likelihood approach that can be expressed via some estimate.

Sample Size Calculation for Selecting a Predictive Marker for a Survival Endpoint

✱Camelia Sima, Memorial Sloan-Kettering Cancer Center, 307 E. 63rd St, 3 Floor, New York, NY 10065, *simac@mskcc.org*; Alexia Iasonos, Memorial Sloan-Kettering Cancer Center; Mithat Gonen, Memorial Sloan-Kettering Cancer Center

Key Words: sample size, power, difference in predictive accuracy, markers, survival endpoint

There is extensive literature covering sample size requirements for survival endpoints in designing prospective clinical trials. This paper provides

methodology for power and sample size calculations for prospectively designing studies that aim to distinguish a more predictive marker from an inferior one, when the endpoint is subject to censoring. The method is applicable to any binary covariate, such as symptom grading (severe/nonsevere), under the assumption that these covariates are measured on the same subjects (paired data). Simulations are performed under alternative schemes by varying the correlation between the covariates; degree of censoring; and parameters of the assumed survival distribution. Power calculations are obtained for a fixed difference in the predictive accuracy between the two covariates, as a function of the sample size. Software is available by the authors.

Multivariate Control Chart for Environmental Monitoring

✱Harry Yang, MedImmune, Inc., One MedImmune Way, Gaithersburg, MD 20878, *yangh@medimmune.com*; Lanju Zhang, MedImmune, Inc.; Iksung Cho, MedImmune, Inc.

Key Words: Environmental monitoring, Control chart, mixtue model

FDA 2004 Guidance and the Code of Federal Regulation governing the manufacture of pharmaceutical products (21 CFR 211) both state the need to establish and follow procedures to prevent microbiological contamination of sterile drug products. Univariate statistical control charts are often used for monitoring either the frequency or magnitude of microbial excursion, in controlled environment. However, there have been no methods in statistical literature dealing with the frequency of zero recovery, and magnitude of non-zero observation simultaneously. In this paper, we propose to use a mixture model to describe microbial recovery from a clean room or isolation system where infrequent and large non-zero observations are indicative of potential contamination. Two statistical tests are suggested, their performances are evaluated through simulations. Also developed are two control charts.

32 Methods in Statistical Genomics ●

ENAR, WNAR, Biometrics Section
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Incorporating Gene Networks into Statistical Tests for Genomic Data via a Spatially Correlated Mixture Model

✱Peng Wei, The University of Minnesota, A460 Mayo, MMC 303, 420 Delaware St SE, Minneapolis, MN 55455, *weixx035@umn.edu*; Wei Pan, The University of Minnesota

Key Words: ChIP-chip, Conditional autoregression, Markov random field, Microarray, Mixture model, Spatial statistics

It is a common task in genomic studies to identify genes satisfying certain conditions, such as differentially expressed genes. Most existing approaches treat genes identically and independently distributed a priori, testing each gene independently. However, it is known that the genes work coordinately as dictated by gene networks. We propose incorporating gene network information into statistical analysis of genomic data. Specifically, we assume that gene-specific prior probabilities are correlated as induced by a gene network: those neighboring genes in the network have similar prior probabilities, reflecting their shared biological functions. We applied both standard mixture model and the proposed method to a real ChIP-chip dataset (and simulated data). The new method was found to be more powerful in discovering the target genes.

Reassessing Shrinkage Estimation in Microarray Experiments Having Within-Array Replicate Spots

Lan Xiao, Michigan State University; ✱Robert Tempelman, Michigan State University, 1205 Anthony Hall, Department of Animal Science, East Lansing, MI 48824-1225, *tempelma@msu.edu*

Key Words: Bayesian, technical replication, false discovery rate, MCMC

Shrinkage based methods have been promising for inference on differential expression in small microarray experiments. Some experiments are characterized by both biological and technical replicates, the latter of which may include several spots per probe on an array. A popular R package, LIMMA, accommodates this issue by assuming a constant correlation for the within-array replicates on each probe, based on the duplicate correlation (dupcor) option. We assert with an example that this assumption may not be true for many experiments and propose a technique that models heterogeneity on within-array correlations between technical replicates across genes. Based on simulation studies, our technique outperformed the dupcor option of LIMMA for ROC and FDR characteristics, particularly when each gene was spotted more than twice on an array and within-array correlations were high.

Bayesian Approach for the Identification of DNA Copy Number Changes in aCGH Data

✱Jie Chen, University of Missouri-Kansas City, 206 Haag Hall, 5100 Rockhill Road, Kansas City, MO 64110, *chenj@umkc.edu*; Ayten Yigiter, University of Missouri-Kansas City; Kuang-Chao Chang, Fu Jen Catholic University

Key Words: Bayesian inferences, change point, uninformative priors, DNA copy numbers, aCGH data

Cancer development or other related diseases are usually relevant to DNA copy number changes on the genome. For the high-throughput array comparative genomic hybridization (aCGH) copy number data, we propose a novel mean and variance change point model (MVCM) for detecting the copy number changes in aCGH data. We use Bayesian approach to study the MVCM for the cases of one change and two changes; and derive the posterior probability for the estimate of the locus or loci of the DNA copy number changes. We carry out simulation studies to evaluate the estimate of the locus (or loci) of the DNA copy number change (or changes) using the derived posterior probability. These simulation results show that the approach is suitable for identifying copy number changes. The approach is successfully applied to the analysis of nine fibroblast cancer cell line data for DNA copy number changes.

A Run-Based Procedure To Identify Time-Lagged Gene Clusters in Microarray Experiments

✱Sunil Mathur, University of Mississippi, Hume 325, Department of Mathematics, University, MS 38677, *skmathur@olemiss.edu*

Key Words: Gene Expression, Microarray, Time-lagged, Cluster

The analysis of gene expression data obtained from microarray experiments can be useful to identify regulatory relationship between genes. Genes with a common functional role have similar expression patterns across different microarray experiments. This similar expression patterns are perhaps due to co-regulation of genes in the same functional group. It is seen that most of the existing methods available for the identification of the regulatory relationships are either made for comparing two genes at a time or methods are not highly efficient in the identification of the regulatory relationships. The procedures adopted by these methods do not use complete information contained in the dataset. In this paper, we propose a statistical procedure which will use the information contained in the dataset to cluster genes which show similar patterns.

A Flexible Semiparametric Test To Detect Associations Between Quantitative Traits and Candidate Genes in Structured Populations with Censored Data

✱Meijuan Li, The University of Minnesota, A460 Mayo Building, 420 Delaware St SE, Division of Biostatistics, Minneapolis, MN 55455, *meijuanl@biostat.umn.edu*

Key Words: Population structure, Censoring, Association mapping, Mixture of polya tree

Several statistical methods for detecting associations between quantitative traits and candidate genes in structured populations have been developed for fully observed phenotypes. However, many experiments are concerned with failure-time phenotypes, which are usually subject to censoring. In this paper, we develop semiparametric statistical methods for detecting association between a censored quantitative trait and candidate genes in the structured population. Our method corrects for population stratification and then models the relationship between trait values, genotypic scores at a candidate marker, and genetic background variables through a semiparametric model, where the error distribution is modeled as a mixture of Polya trees. The proposed method was applied to a real data set of 95 Arabidopsis lines and simulated data to demonstrate power and type I error rate.

Testing for Gene-Environment Interaction: A New Modeling Approach

✱Bin Huang, Cincinnati Children's Hospital Medical Center, Center for Epi/Bio, MLC 5041, 3333 Burnet Ave, Cincinnati, OH 45229, *bin.huang@cchmc.org*; Altaye Mekibib, Cincinnati Children's Hospital Medical Center; Siva Sivaganisan, University Cincinnati

Key Words: Gene-Environmetal Interaction, 8 fold table, Relative Risk, Case-Control Study, Cohort Study

The Gene-Environment (GE) interaction refers to "a different effect of an environmental exposure (genotype) on disease risk in persons with different genotypes (environmental exposure)" (Ottman 96). Estimating GE interaction in a case-control study setting has been a subject of many resent studies. Existing methods often assume rare disease or gene-environmental independence assumption. Logistic regression or log linear modeling were used, and the coefficient of the interaction term is used as the effect measure for GE interaction. In this study, we propose a new modeling approach. We demonstrated that under the rare disease and/or independence assumption, our approach is consistent with the existing methods, while these assumptions are not necessary. Simulation studies as well as case studies will be conducted to assess biase and efficiency.

Power Consideration for Incomplete Data Sets in Genetic Studies

✱Hemant K. Tiwari, The University of Alabama at Birmingham, AL 35294, *HTiwari@ms.soph.uab.edu*; Miguel Padilla, Old Dominion University; Howard Wiener, The University of Alabama at Birmingham; Jerome Reiter, Duke University

Key Words: Missing Data, Genetics, Power

Missing data are ubiquitous and inevitable in any area of research. Potential problems due to missing data include bias in parameter estimation, inflated type I error rates, loss of power, and often complexity of executing the analysis. Determining the appropriate analytic approach in the presence of incomplete data has been an active area of research in recent decades. Although remarkable progress has been made in the development of statistical approaches that provide valid and reliable parameter estimates in the presence of missing data, missing data problems in genetic studies with multivariate phenotypes and large numbers of available genotypes typed on individuals have not been adequately addressed.

33 Likelihood-Based Methodology

Biometrics Section, Section on Nonparametric Statistics, IMS

Sunday, August 3, 2:00 p.m.–3:50 p.m.

Analyzing Incomplete Data Subject to a Threshold Using Empirical Likelihood Methods: An Application to a Pneumonia Risk Study in an ICU Setting

✱Jihnhee Yu, University at Buffalo, NY 14214-3000, *jinheeyu@buffalo.edu*; Albert Vexler, National Institute of Child Health and Human Development; Lili Tian, University at Buffalo

Key Words: Empirical likelihood ratio test, Incomplete data, Ventilator associated pneumonia, Two-sample test, Linear regression

The first detection of pneumonia for inpatients at an intensive care unit (ICU) needs external symptom evaluation by the clinical pulmonary infection score (CPIS). When the CPIS is above a threshold value, bronchoalveolar lavage (BAL) is performed to count actual bacterial pathogens. Thus, CPIS and BAL results are closely related and important indicators of pneumonia while BAL data are incomplete. To compare the pneumonia risks between treatment groups for such incomplete data, we derive a new method that combines empirical likelihood ratio test with classical testing utilizing parametric models. The asymptotic Type I error of the proposed method is investigated theoretically. Monte-Carlo study confirms that the method has good power properties. The method is applied to the data obtained in clinical practice setting and effectively compares pneumonia risks between treatment groups.

Empirical Likelihood Inference for Censored Median Regression Model via Nonparametric Kernel Estimation

✱Yichuan Zhao, Georgia State University, Department of Mathematics and Statistics, Atlanta, GA 30303, *matyiz@langate.gsu.edu*; Feiming Chen, Spectra Marketing Systems

Key Words: Coverage probability, Least absolute deviations, Right censoring, Confidence regions, Empirical likelihood, Missing information principle

An alternative to the accelerated failure time model is to regress the median of the failure time on the covariates. In the recent years, censored median regression models have been shown to be useful for analyzing a variety of censored survival data with the robustness property. Based on missing information principle, a semiparametric inference procedure for regression parameter has been developed when censoring variable depends on continuous covariate. We apply an empirical likelihood ratio method (EL) to the model and derive the limiting distributions of the estimated and adjusted empirical likelihood ratios for the vector of regression parameter. The EL confidence regions for the unknown vector of regression parameters are obtained accordingly. The simulation results suggest that the EL methods outperform the normal approximation based method in terms of coverage probability.

A Simulation Study Comparing Likelihood and Nonlikelihood Approaches in Analyzing Overdispersed Count Data

✱Stanley Xu, Kaiser Permanente Colorado, 6081 S Paris St, Greenwood Village, CO 80111, *stan.xu@kp.org*; Gary K. Grunwald, University of Colorado; Richard Jones, University of Colorado

Key Words: Poisson mixture, Overdispersion, Gamma, Lognormal, Inverse Gaussian, Maximum likehood estimate

In this simulation study, overdispersed count data were simulated under the Poisson mixtures, then analyzed with likelihood approaches including three Poisson mixtures with gamma, lognormal, and inverse Gaussian distributions and non-likelihood approaches including the robust sandwich estimator and quasilikelihood. Our results indicated that 1) when the count data are mildly overdispersed, there are virtually no differences in type I error rate, standard error of the main effect, and empirical power among the five methods; 2) when the count data are very overdispersed, none of these five approaches is robust to model misspecification as evaluated by type I error rate, standard error of the main effect, and empirical power. We recommend choosing the one with the smaller -2 log likelihood value between the Poisson mixtures with the gamma distribution and with the lognormal distribution.

Empirical Likelihood Test for Equality of Means of Populations Containing Many Zeros

✱Le Kang, State University of New York at Buffalo, 253 Farber Hall, 3435 Main St, Buffalo, NY 14214, *lekang@buffalo.edu*; Albert Vexler, The State University of New York at Buffalo; Lili Tian, University at Buffalo; Germaine B. Louis, National Institute of Child Health and Human Development

Key Words: mixture model, parametric likelihood, empirical likelihood, likelihood ratio, monte carlo, type I error

Endometriosis is a complex disease occurring in menstruating women. Recent study suggested organochlorine pesticides such as Aldrin and beta-BHC, might have important differences in their concentrations in serum between women with and without disease. The relevant dataset includes zero observations that were recorded due to the measurement problems, e.g., in part related to instrument precision [Schisterman et al.(2006)]. Instead of assuming a parametric mixture model with lognormal distribution [Zhou et al.(1999)], we assume that biomarkers Aldrin and beta-BHC follow a nonparametric mixture model and extend the existing empirical likelihood techniques, which are based on one population [Chen et al.(2003)], to test the equality between two population means. Our broad Monte Carlo study compares the proposed test to some parametric approaches with respect to the Type I error controls.

Summarizing Likelihood Functions as Functions

✱Michael B. Brimacombe, New Jersey Medical School, 185 S. Orange Ave (MSB F-633), Newark, NJ 07101-1709, *brimacmb@umdnj.edu*

Key Words: likelihood, meta-analysis, functional data

In a growing number of meta-analytic settings information is available to statisticians from a set of existing studies. The goal of the analysis is to achieve an overall statistical assessment of the treatment under investigation. From a statistical perspective, the likelihood function for each study, if available, summarizes all relevant information regarding the parameter of interest. Viewing the likelihood function as a function of the parameter of interest and placing the likelihood into graphical and random effects setting allows for the use of functional data interpretations to derive summary pseudo-likelihoods that can be used to provide summary information for the entire set of studies in question. This can be used to generate pseudo likelihood ratio tests to examine patterns in the overall dataset. Examples are drawn from the biomedical literature.

Conventional Inference via a Holistic Approach

✱Michael P. Jones, Macquarie University, C3A 520, North Ryde, International 2109 Australia, *mike.jones@psy.mq.edu.au*

Key Words: multivariate methods, logistic regression, analytic approach

Aim: Compare conventional statistical inference with a more holistic approach. Method: Using a study of gastrointestinal diseases as a vehicle we address the practical research question “Are dyspepsia symptoms associated with weight loss ?” Conventional inference uses unconditional logistic regression to isolate statistically independent influences on weight loss. In the alternate approach factor analysis is used to quantify latent symptom dimensions on which individuals are then formed into mutually exclusive clusters using a non-hierarchical algorithm and the cluster symptom profiles described. Weight loss prevalence is then overlaid on cluster profiles. Findings: While some conclusions would be arrived at either way, the holistic approach allowed insights hidden by conventional inference. Communication with nonstatisticians is easier and more effective.

Testing the Homogeneity of the Means of Several Groups of Count Data in the Presence of Unequal Dispersions

✱Krishna Saha, Central Connecticut State University, Department of Mathematical Sciences, 1615 Stanley Street, New Britain, 06050, *sahakrk@ccsu.edu*

Key Words: Score test, Double extended quasi-likelihood, Homogeneity of the means, Negative binomial model, Dispersion, Power

Extra-dispersion is a common phenomenon in the biostatistical field when count data exhibit extra-variation relatively a Poisson model. This arises when the data are grouped or when the assumption of independence is violated. In this paper, procedure for testing the equality of the means of several groups of count data when extra-dispersions among the groups are unequal is developed based on the adjusted count using the concept of the design effect and size effect proposed by Rao and Scott (1999, Statistics in Medicine). We also obtain the score type test statistics using the quasi-likelihoods based on the mean-variance structure of the negative binomial, and study the properties and performance characteristics of these statistics. The simulation results indicate that the statistic based on the adjusted count data holds the best performance characteristics over the other statistics.

34 Issues Related to the Design of Clinical Trials ●▲

Biopharmaceutical Section, Biometrics Section
Sunday, August 3, 2:00 p.m.–3:50 p.m.

Optimal Ray Design for Drug Combinations

✱Yuehui Wu, GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, PA 19426, *yuehui.2.wu@gsk.com*; John J. Peterson, GlaxoSmithKline Pharmaceuticals

Key Words: drug combination, fixed-ratio ray design, locally D-optimal design

When two drugs are studied as a mixture in the drug development, one expects the combination works better than any of the drugs works alone. To investigate this, we consider the fixed-ratio ray design that the ratio of drug A to drug B is fixed on a single ray for different doses and the observation will be collected on a few rays (ratios) in the whole experiment. The model under consideration is the nonlinear logistic model and the locally D-optimal design techniques will be applied to locate the optimal ray design which answers the following questions: how many rays are needed? What ratios should one chose? How many subjects should be located on each dose? We also consider the D-optimal design under some restrictions such as the total number of rays and the number of different doses on each ray is limited. Comparison of the optimal design to a random selected ray design is also made.

Optimal Designs for Toxicity Function with a Downturn

✱Seung Won Hyun, University of Missouri-Columbia, *swhrg7@mizzou.edu*; Nancy Flournoy, University of Missouri

Key Words: Dose-finding, Experimental design, Toxicological assays, Receptor-meditated response, A-optimality for adjacent treatment comparisons, D-optimality

In many toxicological assays, interactions between primary and secondary effects may cause a downturn in the response functions at high concentrations. Downturns have been studied by Margolin (1981), Bretz and Hothorn (2001), Welshons et al. (2003) and others. We obtain optimal designs for several models of such nonmonotone concentration response functions. A related situation occurs in clinical trials on which success is measured as a function of both toxicity and safety. See, for example, Dragalin, Fedorov and Wu (2006) and Rabie and Flournoy (2004).

Orthogonal Experimental Designs Can Disentangle Confounding in Database Studies

✱Charlie H. Goldsmith, McMaster University, Biostatistics Unit, Martha H322, St Jospeh's Healthcare Hamilton, 50 Charlton Ave E, Hamilton, ON L8N 4A6 Canada, *goldsmit@mcmaster.ca*; Lehana Thabane, McMaster University; Gary Foster, St. Joseph's Healthcare Hamilton; Eleanor M. Pullenayegum, St. Joseph's Healthcare

Key Words: Orthogonal designs, confounding, databases, bootstrapping

To determine whether confounding in databases studies can be disentangled using orthogonal factorial designs. Factorial designs permit estimation of interactions and main effects. Replicates of cases were sampled according to a bootstrap scheme to create a factorial design, which was then used to estimate the impact of main effects and interactions. All estimable main effects and interactions are combined like imputation to provide an analysis of the factors and interactions better than with the analysis of the database because of confounding. Ideas will be shown using an osteoporosis database, a 2**3 complete factorial, 10 replicates and bootstrapping. A comparison with the usual analysis, feasibility and interpretations of these bootstrapped effects will be shown. Orthogonal designs created via bootstrapping can disentangle interactions between factors in database studies.

Optimization of Design Features in Clinical Trials with a Sensitive Subgroup

✱Yan D. Zhao, Eli Lilly and Company, Lilly Corporate Center, DC 6054, Indianapolis, IN 46285, *zhaoyan1@gmail.com*; Alex Dmitrienko, Eli Lilly and Company; Roy Tamura, Eli Lilly and Company

Key Words: Clinical trial design, sensitive subgroup, optimal design, type I error control, power, simulation

This talk introduces a method for optimizing the design of clinical trials with a sensitive subgroup of patients (i.e., a subgroup that is more likely to benefit from the treatment than the overall population). Given a sensitive subgroup (defined by a classifier [e.g., a clinical marker or pharmacogenomic marker]), the study's outcome is declared positive if the treatment effect is established in the overall population or the subgroup. The proposed method aims at achieving an optimal balance between the power of the overall and subgroup tests under restrictions that reflect the trial's objective. The restrictions include control of the overall Type I error rate and requirement to show that the overall treatment effect, if it is demonstrated, is not limited only to the sensitive subgroup. Operating characteristics of the proposed design are assessed via a simulation study.

Strengthening the Understanding of the Relationship Between Survival Designs and Test Statistics

✱Jitendra Ganju, Amgen, Inc, 1120 Veterans Blvd, South San Francisco, CA 94080, *jganju@amgen.com*; Guoguang (Julie) Ma, Amgen, Inc

Key Words: logrank statistic, follow-up time

In the planning of survival trials, the role of patient follow-up time introduces a level of complexity not encountered in other (nonsurvival) trials. Of two commonly used survival designs, one fixes the patient follow-up time whereas the other allows time to vary. When the follow-up time is fixed the number of events varies. Follow-up time varies when the number of events is fixed. We explain how the two designs time to event (e.g., logrank) and binomial (e.g., relative risk, risk difference) test statistics.

Design and Interim Monitoring of a St. Jude Clinical Trial Based on the Gompertz Distribution

✱Arzu Onar, St. Jude Children's Research Hospital, 332 North Lauderdale St, Mail Stop 768, Memphis, TN 38105, *arzu.onar@stjude.org*; James M. Boyett, St. Jude Children's Research Hospital; Robert P. Sanders, St. Jude Children's Research Hospital; Amar Gajjar, St. Jude Children's Research Hospital

Key Words: Clinical Trial, Gompertz Distribution, Sample size estimation

This talk details the statistical issues encountered in a recent single-arm trial at St Jude Children's Research Hospital for young children with embryonic brain tumors. The trial design aimed to estimate EFS distribution, which has historically been characterized by a rapid decline followed by an asymptote, representing patients who achieve long-term disease control. This pattern makes the use of conventional exponential-based designs inappropriate. So instead we used the Gompertz distribution which can accommodate increasing, decreasing and constant hazards. The parameters are easy to estimate and have intuitive interpretations. Further the log-rank type procedures can be used for testing. Sample size calculations were based on estimating the cure rate fraction and a single interim analysis was proposed. The latter however poses some additional challenges which will be detailed.

Noninferiority Trial Designs for Binomial Rate Differences and Odds Ratios

✱Joan Hilton, University of California, San Francisco, Department of Epidemiology & Biostatistics, 185 Berry Street, Suite 5700, San Francisco, CA 94107-1762, *joan@biostat.ucsf.edu*

Key Words: noninferiority, optimal allocation, design parameters, sample size calculation, reparameterization, binomial response

In the setting of binomial rates that differ under H0 and are equal under HA, we show that the marginal response rate and the noninferiority margin are a convenient "design pair" for noninferiority trials. Further, the minimum overall sample size, N, and optimal allocation ratio associated with fixed error rates depend on how the margin is parameterized. Since investigators commonly use the difference between test and control response rates (Delta) for design and the log-odds of response (logPsi) for analysis, we quantify effects on power of switching parameterizations of the margin. We also model the ratio N_Delta/N_Psi as a function of a wide range of design pairs; the regression estimates can be used at the design stage to identify pairs for which the margin's parameterization should not be interchanged between design and analysis. Last, we propose ways to quantify the marginal rate.

35 Introductory Overview Lecture: Interdisciplinary Communications: Functional Data Analysis and Differential Equation Models

ASA, ENAR, IMS, SSC, WNAR

Sunday, August 3, 4:00 p.m.–5:50 p.m.

From Functional Data Analysis to Differential Equation Models, with Applications

✱James O. Ramsay, McGill University, 2748 Howe St., Ottawa, ON K2B 6W9, *ramsay@psych.mcgill.ca*

Dynamical models explain how a system responds to a change in the system itself or in its environment. They are usually expressed as a set of differential equations that explicitly link change as defined by time derivatives to the state of the system and its inputs. Nonlinear systems are especially interesting because of their potential to capture complex system behavior by a seemingly simple set of equations. A new method will be outlined which permits the estimation of nonlinear system parameters from noisy data along with confidence regions and other inferential methodology. The procedure will be illustrated with examples drawn from chemical engineering.

Statistical Methods for Inverse Problems of Differential Equations, with Biomedical Applications

✱Hulin Wu, University of Rochester, Department of Biostatistics and Computational Biology, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, *hwu@bst.rochester.edu*

Differential equation models are widely used to describe dynamic systems in many scientific fields such as engineering, physics, econometrics, and biomedical sciences. It is challenging to develop statistical methods to estimate the parameters in a nonlinear differential equation (NDE) model when the closed-form solution is not available. Model identifiability of the NDE models will be discussed and statistical estimation methods will be introduced and reviewed for both constant and time-varying parameter estimation in NDE models. Biomedical application examples will be used to illustrate the proposed methods. A user-friendly software package, DEDiscover, will be introduced.

36 Analyzing Genetic Data from Health Surveys: Opportunities and Challenges ●

ENAR, Biopharmaceutical Section, WNAR, Section on Health Policy Statistics, Section on Statistics in Epidemiology, Biometrics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Genetic Research in the National Health and Nutrition Examination Surveys

✱Christopher Sanders, National Center for Health Statistics, Hyattsville, MD 20782, *bxu5@cdc.gov*

Key Words: NHANES, genetics, genome-wide association study, surveys, official statistics

NHANES is a series of nationally representative health surveys run by the National Center for Health Statistics. In the past two years we have collected genetic data in over a hundred genes of from ~7,100 participants in NHANES III. The analyses of these genetic variations linked to thousands of NHANES lab, questionnaire, and environmental variables has led researchers to produce prevalence estimates, explore genotype-phenotype associations, and create new research methods to analyze genetic data within the framework of the NHANES survey design. Initiatives are in progress for the collection of one million genetic variations in ~15,000 participants in the NHANES III and NHANES 1999-2002 DNA banks. This discussion will provide an overview of the ongoing genetic research that has been done using NHANES data and highlight the potential use of this data in future research.

Estimating Family Relationships Using DNA Fingerprints Within the NHANES-III Household Survey

✱Hormuzd A. Katki, National Cancer Institute, 6120 Executive Blvd Room 8016 MSC 7244, Rockville, MD 20852-4910, *katkih@mail.nih.gov*

Key Words: NHANES, genetics, demography, surveys, official statistics

The NHANES-III survey did not accurately assess familial relationships within household members. We estimated family relationships using each participant's Identifiler(tm) DNA fingerprint, a set of 15 DNA markers that overlap with the markers used by the FBI for forensic identification. We adapted two methods of estimating familial relatedness (an exact method requiring allele frequencies and a less efficient method that does not require allele frequencies) to survey data, accounting for genotyping error rates and potential correlation of alleles within ethnic groups. We both tested and estimated family relationships for all pairs of potential relatives within a household. This work will facilitate future family-based association and linkage studies within NHANES-III.

Testing for Hardy-Weinberg Equilibrium and for Disequilibrium Across Population Strata in Complex Survey Sample

✱Yan Li, National Cancer Institute, 20852-4910, *lisherry@mail.nih.gov*

Key Words: NHANES, genetics, genome-wide association study, surveys, official statistics

In population genetics, the use of representative random samples of the target population can avoid ascertainment bias. Genetic data from over a hundred genes were collected in NHANES III. Surveys such as the NHANES have complex sample designs with weighting that can inflate variances and alter the expectations of test statistics. Thus, classical tests of Hardy-Weinberg Equilibrium (HWE) and Homogeneity of Hardy-Weinberg Disequilibrium (HHWD) for simple random samples are not suitable for complex samples. We propose using Wald tests for HWE and generalized score tests for HHWD, modified for complex samples. Simulation studies are used to investigate the finite sample properties of the proposed tests. Rao-Scott corrections were found to improve test type I error properties. Our methods are applied to the NHANES III genetic data for three alleles involved in metabolizing lead.

37 Spatio-Temporal Dynamic Models ●▲

Section on Bayesian Statistical Science, Section on Statisticians in Defense and National Security, Section on Physical and Engineering Sciences, Section on Statistics and the Environment

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Space-Time Random-Effects Models

✱Noel Cressie, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Room 404, Columbus, OH 43210-1247, *ncressie@stat.osu.edu*; Lei (Emily) Kang, The Ohio State University

Key Words: Bayesian prediction, fixed rank kriging, hierarchical models, Kalman filtering, mixed-effects models, posterior distribution

In this talk, we propose the use of space-time random-effects (STRE) models and space-time mixed-effects (STME) models to fit to environmental data distributed in space and time. Both models can be viewed as a hierarchical statistical model that includes a data model and a (space-time) process model. We consider massive datasets and hence a dimension-reducing STME. Unknown parameters in the model can be estimated or a prior distribution (parameter model) can be posited. In the former case, an empirical Bayes procedure yields a space-time Kalman Filter that we call the Fixed Rank Kalman Filter (FRKF). In the latter case, an MCMC algorithm is needed for optimal filtering, which exploits the dimension reduction referred to earlier. Our methodology is applied to NASA's Multi-angle Imaging SpectroRadiometer (MISR) Aerosol Optical Depth data.

Combining Computer Model Output and Observational Data: Forecasting Climatic and Weather-Related Processes

✱Mark Berliner, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210, *mb@stat.osu.edu*

Key Words: Climate change, Hierarchical Bayesian analysis, Experimental design

Modern science relies on both computer modeling and continuing improvements in observational assets. However, both of these information sources are subject to error. Hence, combining these sources in a fashion that reflects their uncertainties is critical. The approach suggested is a hierarchical Bayesian formulation involving the statistical modeling of computer model output. The approach also enables direct formulation of experimental design for combining computer models and observations. The main notions are illustrated in an extended example involving climate and climate change analyses.

Tracking Space Junk

✱Dave Higdon, Los Alamos National Laboratory, *dhigdon@lanl.gov*; Earl Lawrence, Los Alamos National Laboratory

Key Words: filtering, Markov chain Monte Carlo, model selection

There are currently thousands of resident space objects orbiting the planet, and this number will grow substantially over the coming years. The challenge is to track these objects with the help of physical observations as well as computational models for their trajectories. This task is made difficult by limited observational resources, and by external factors such as atmospheric conditions and space weather that can alter an object's trajectory. In this talk we develop a Bayesian modeling approach that combines these deterministic simulation models with observations to track multiple objects.

38 Semiparametric Modeling of Nonlinear and Nonstationary Time Series ●

Business and Economics Statistics Section, IMS

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Consistent Nonparametric Tests on Parametric Smooth Coefficient Models

✱Yiguo Sun, University of Guelph, Department of Economics, Guelph, ON N1G 2W1 Canada, *yisun@uoguelph.ca*; Zongwu Cai, The University of North Carolina at Charlotte; Qi Li, Texas A&M University

Key Words: Nonstationary, nonlinearity, semiparametric estimation, Degenerate U-Statistic, Central limit theorem, Degenerate U-Statistic, Central limit theorem

In this paper, we consider the problem of consistent model specification testing with nonstationary data. We propose a simple kernel-based nonparametric test for testing the null hypothesis of constant coefficients against nonparametric smooth coefficients in a varying coefficient model with some of the covariates being nonstationary data. We establish new central limit theorems for generate U-statistics with nonstationary data and use them to derive the asymptotic distributions of the proposed test statistic under both null and alternative hypotheses. A Monte Carlo simulation is conducted to illustrate the finite sample performance of the proposed test statistic.

Spline-Backfitted Kernel Smoothing of Additive Coefficient Model

✱Rong Liu, Michigan State University, Department of Statistics and Probability, A413 Wells Hall, East Lansing, MI 48824, *liurong@stt.msu.edu*; Lijian Yang, Michigan State University

Key Words: bandwidths, kernel, knots, local linear estimator, mixing, nonparametric regression

Additive coefficient model (Xue and Yang 2006a,b) is a flexible tool for multivariate regression and time series analysis that circumvents the "curse of dimensionality." We propose spline-backfitted kernel (SBK) and spline-backfitted local linear (SBLL) estimators for the component functions in the additive coefficient model that is both (i) computationally expedient so it is usable for analyzing very high-dimensional data, and (ii) theoretically reliable so inference can be made on the component functions with confidence. In addition, it is (iii) intuitively appealing so it does not intimidate practitioners. Simulation experiments have provided strong evidence that corroborates with the asymptotic theory. The SBLL procedure is applied to a varying coefficient extension of the Cobb-Douglas model for the US GDP that allows non-neutral effects of the R&D on capital, labor and TFP.

Empirical Likelihood-Based Inference for Nonlinear Diffusions

✱Ke-Li Xu, University of Alberta, 3-40N Business Building, School of Business, Edmonton, AB T6G 2R6 Canada, *keli.xu@ualberta.ca*

Key Words: diffusion models, empirical likelihood, local linear smoothing, nonparametric modeling

This paper provides a new approach to constructing point-wise confidence intervals for nonparametric drift and diffusion functions in a continuous-time diffusion model via empirical likelihood (EL) in conjunction with local linear smoothing. Limit theories are developed for the EL ratios for both drift and diffusion functions by means of increasing time span and shrinking observational intervals. Our results apply to both stationary and nonstationary recurrent diffusion processes. Compared to the asymptotic normality-based method, our approach avoids the difficulty of variance estimation by choosing the variance estimator implicitly, and it has the additional advantage of automatic determination of the shape of the resultant confidence interval. Simulation studies show that for both the drift and diffusion functions the EL confidence intervals work remarkably well.

39 Post-Imputation Variance Estimation ●

Section on Survey Research Methods, Section on Nonparametric Statistics, Section on Government Statistics, Social Statistics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Multiple Semiparametric Imputation

✱David R. Judkins, Westat, 1650 Research Blvd., Rockville, MD 20850, *DavidJudkins@westat.com*; Andrea R. Piesse, Westat; Tom Krenzke, Westat

Key Words: Post-imputation variance estimation, Hot-deck imputation

In 2007, Judkins, Krenzke, Piesse, Fan, and Haung reported on the performance of a new semi-parametric imputation algorithm designed to impute entire questionnaires with minimal human supervision while preserving important first- and second-order distributional properties. In the 2008 paper, we will report on procedures for post-imputation variance estimation to be used in conjunction with the semi-parametric imputation algorithm. The basic idea will be to weaken the quality of donor-recipient matching adequately to enable multiple imputations that are near enough to proper imputations in order for Rubin's formula to be reasonably employed. The presentation will include results from a simulation study.

Multiply Imputing Potential Outcomes To Estimate Individual Causal Effects

✱Susanne Rässler, Otto-Friedrich University Bamberg, Feldkirchenstr. 21, Bamberg, 96045 Germany, *susanne.raessler@uni-bamberg.de*; Donald B. Rubin, Harvard University

Key Words: Matching, Rubin's Causal Model, evaluation, labor market policy

The German Federal Employment Agency is using a comprehensive evaluation system to measure the efficiency of the large variety of job-training programs. Because the assignment to any of the possible training programs (i.e., treatments) typically is not based on a random process, corrections have to be done before drawing causal inferences. Basically, we try to answer the question for each job-training program and each trained person, if a person had not been job trained, how much time would have passed until the person found employment? Based on Rubin's Causal Model we multiply impute these missing potential outcomes to get estimates of the individual causal effects. Since this is a massive imputation task, innovative and complex algorithms as well as an automation of the whole process is required. In this talk, we describe our approach, the imputation routines, and present some results.

Fractional Imputation Using a Parametric Imputation Model

✱Jae-kwang Kim, Yonsei University, Republic of Korea, *kimj@yonsei.ac.kr*; Wayne Fuller, Iowa State University

Key Words: Missing data, EM algorithm, Importance sampling, Item nonresponse

Under item nonresponse, imputed values are often generated from a parametric model. When the imputation model belongs to a parametric family of distributions, imputation requires estimated parameter values in order to impute for the missing data. We propose parametric fractional imputation, which is a parametric approach for generating imputed values. The proposed imputation method provides very efficient parameter estimates for the parameters specified in the model and, at the same time, also provides reasonable estimates for parameters that were not considered in the imputation model, such as domain means. Thus, the proposed imputation method is a very useful tool for general-purpose data analysis. Variance estimation is covered and results from a limitd simulation study are presented.

40 Statistical Methods for Multivariate Dental Data ●▲

Biometrics Section, Biopharmaceutical Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Bayesian Modeling of Multivariate Spatial Binary Data: Applications to Dental Epidemiology

✱Brian Reich, North Carolina State University, 2501 Founders Drive, 209H Patterson Hall, Box 8203, Raleigh, NC 27695, *reich@stat.ncsu.edu*; Dipankar Bandyopadhyay, Medical University of South Carolina; Elizabeth Slate, Medical University of South Carolina

Key Words: dental data, spatial modelling, Bayesian methods, autologistic prior

A clinical study was conducted at the Medical University of South Carolina (MUSC) to determine the dental health status of Type-2 diabetic Gullah African-American adults (residing in coastal sea islands of South Carolina) who have a minimal Caucasian genetic admixture. We develop a multivariate model for this spatial binary data (binary responses indicative of whether a tooth surface be either decayed, missing or filled, clustered across subjects and across tooth) using a random effects autologistic regression model that controls for the correlation within tooth surfaces and spatial correlation among neighbor teeth. We compare this autologistic model with covariates with other possible models to see the improvement in predictions, assess spatial associations and also to determine the effects of covariates on the dental caries status of this understudied population.

Bayesian Generalized Latent Variable Models for Spatial Correlated Binary Data with Applications to Dental Outcomes

✱David Todem, Michigan State University, Division of Biostatistics, Department of Epidemiology, B601 West Fee Hall, East Lansing, MI 48824, *todem@msu.edu*; Yanwei Zhang, Michigan State University; KyungMann Kim, University of Wisconsin-Madison; Emmanuel Lesaffre, Erasmus University

Key Words: Spatial Correlation, Undirected Graphical Gaussian model, Nested Design, Multivariate Conditional Autoregressive Models, Latent Variable

Analysis of dental caries is traditionally based on DMFS and DMFT scores, which are a summary of a caries experience for each individual. Although this approach has aided our understanding of the pattern of dental caries, there are still some fundamental questions that remain unanswered. As an example, it is well believed among dentists that there are spatial symmetries in the mouth with respect to caries, but this has never been shown in a statistical sense. An answer to this question requires the analysis to be performed at the tooth level. This then necessitates the use of methods for correlated data. In this paper, we propose a Bayesian generalized latent variable model for the incidence of dental caries at each location while taking into account the unique spatial co-dependence of teeth within a given mouth. Data from a cross-sectional survey are used to illustrate the methodology.

Analysis of Nested Observations Over Teeth and Time: A Combined Estimating Equations Approach

✱Julie A. Stoner, University of Oklahoma Health Sciences Center, 801 NE 13th Street, CHB 309, PO Box 26901, Oklahoma City, OK 73190, *julie.stoner@yahoo.com*; Brian G. Leroux, University of Washington

Key Words: Correlated Data, Dental Data, Generalized Estimating Equations, Longitudinal Data, Nested Data

Dental research data are often nested within a given subject's mouth and repeated measures are made longitudinally. Cluster sizes are large, relative to the number of independent patients, and correlation structures are complicated. A marginal regression modeling method that optimally weights and combines between-cluster and within-cluster sources of information, including contrasts among nested observations within a cluster, among longitudinal measures, and the interaction among nested and longitudinal measures, is presented. Method performance is studied by simulation and an application to a periodontal trial. The method results in increased estimation efficiency, particularly when the degree of correlation is not constant across clusters, while maintaining stability and coverage probabilities and avoids direct correlation parameter estimation. (NIH grants R01DE15651, R01DE012872)

Flexible Accelerated Failure Time Frailty Models for Multivariate Interval-Censored Data with an Application in Caries Research

✱Emmanuel Lesaffre, Erasmus University, Dr. Molewaterplein 50, Department of Biostatistics, Rotterdam, 3000 The Netherlands, *e.lesaffre@erasmusmc.nl*; Arnost Komarek, Charles University in Prague

Key Words: Clustered data, Density smoothing, Markov Chain Monte Carlo, Survival analysis

In this paper we consider the relationship of covariates to the time to caries of permanent first molars. This involves an analysis of multivariate doubly-interval-censored data. To describe this relationship we suggest an accelerated failure time model with random effects taking into account that the observations are clustered. Indeed, up to four permanent molars per child enter into the analysis implying up to four caries times for each child. Each distributional part of the model is specified in a flexible way as a penalized Gaussian mixture (PGM) with an overspecified number of mixture components. A Bayesian approach with the MCMC methodology is used to estimate the model parameters and a software package in the language has been written that implements it.

41 Communicating with Nonstatisticians ●▲

Section on Quality and Productivity, Section on Physical and Engineering Sciences, Section on Bayesian Statistical Science, Social Statistics Section, Section on Government Statistics, Section on Health Policy Statistics

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Statistics and Due Process: Two Statistically Significant Issues in Law

✱Mark G. Haug, The University of Kansas School of Business, 1300 Sunnyside Ave., Lawrence, KS 66045, *mhaug@ku.edu*

Key Words: law, evidence, epidemiology

Mark Haug, PhD, JD, will discuss two significant legal issues concerning evidence that determine the due process rights of thousands of litigants each year, and why statistical and epidemiological education are essential for the proper resolution of these evidentiary questions. The two issues are 1) the statistical evaluation of expert evidence and whether it is admissible at trial, and 2) how statistical evidence, once admitted, can be misused at trial.

Providing Biostatistical Support to an Understaffed and Underfunded Region

✱Liam O'Brien, Colby College, Department of Mathematics, 5838 Mayflower Hill, Waterville, ME 04901, *lobrien@colby.edu*

Key Words: statistical consultation, statistical applications, undergraduate education

I will discuss my experiences providing biostatistical support to community- and state-based researchers in central Maine. The region provides an applied statistician such as myself with a wide variety of interesting data and many opportunities to get involved in all aspects of research. While all the people I encounter are impressively driven, capable, and focused on making the rural region in which we live a better place, they have very different backgrounds and each project presents its own set of communication challenges in addition to statistical challenges. These issues have proven to be a boon for getting undergraduate students involved in real-world research in all aspects of statistical science.

Communicating Statistics to Engineers and Managers

✱Gerald Hahn, GE Corporate Research and Development (retired), Applied Statistics, 1404 Orlyn Drive, Schenectady, NY 12309, *gerryhahn@yahoo.com*

Key Words: Effective communication, Effective Presentation, Industry, Business, Elevator speech

How can we get across the value of statistics to engineers and managers (some of whom may have been turned off by an introductory statistics course in school)? How can we convince others of the need for our involvement at the definition and data gathering stage of projects? How can we communicate effectively during the course of our work and in presenting our findings? How can we summarize our past year contributions in a 20-minute presentation to our division Vice-President or company President? In this talk, I will try to address the preceding questions, based upon 46 years of experience in the trenches.

42 Visualizing Large Data Sets ●▲

Section on Statistical Graphics, Section on Nonparametric Statistics, Section on Statisticians in Defense and National Security, Section on Physical and Engineering Sciences, Section on Government Statistics
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Visualization and Analysis of Streaming Data

✱Simon Urbanek, AT&T Labs - Research, 180 Park Ave, Florham Park, NJ 07932, *urbanek@research.att.com*

Key Words: visualization, large data, streams, interactive graphics

Streaming data impose constraints on the methods which can be used for their visualization and analysis. New methods and techniques need to develop to deal with the transience and rate of accumulation as well as the size of the data. In this talk we will show methods for visualization and their implementation that deal with those problems.

Ensembles of Models for Understanding Large Data

✱Hadley Wickham, Iowa State University, Ames, IA 50010, *h.wickham@gmail.com*

Key Words: modeling, graphics, GGobi, R

In this talk, I will explore the use of large numbers of simple models to aid the understanding of large, complex data sets. Models ensembles are commonly used to enhance predictive ability, but can they also enhance our understanding of the underlying data? I will illustrate these ideas with R and Ggobi.

Scagnostics and Related Techniques for Automodeling-Based Visualization

Leland Wilkinson, Northwestern University; ✱Graham Wills, SPSS, Chicago, IL, *gwills@spss.com*

Key Words: visualization, statistical graphics

When faced with a data set that is both wide and long, it can be a challenge to see even a simple overview of the data. Validating data, transforming it and understanding simple univariate structures is both tedious and error-prone. With hundreds or thousands of variables, when we move beyond univariate analysis to multivariate analysis, some form of automatic aid becomes necessary. In this talk we introduce Tukey and Tukey scagnostics and develop graph theoretic methods for implementing their procedure on large datasets. We describe a prototype automatic modeling system (for a very loose definition of 'modeling') and explore its advantages and deficiencies for the automatic modeling of large data sets.

43 New Statistical Methods for Genomewide Association Studies ●▲

IMS, Biopharmaceutical Section, WNAR, Section on Statistics in Epidemiology
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Opportunities and Challenges of Whole-Genome Association Studies

✱David Siegmund, Stanford University, Sequoia Hall, Serra Mall, Stanford, CA 94305, *dos@stat.stanford.edu*

Key Words: association, genome, linkage

A unified framework for association studies and for linkage studies will be described. It will be shown that although one expects association studies to have substantially more power than linkage studies, standard statistics to test for association are less robust to modeling errors, so there can be a significant loss of power if one fails to include dominance or gene-gene interaction in the model. For approximating significance levels, simple improvements on the Bonferroni bound and on the central limit theorem will be discussed.

Extensions of Conditional Likelihood Bias Adjustment for Disease Association Risk Estimates in Whole-Genome Scans

✱Fred Wright, The University of North Carolina at Chapel Hill, 3107B McGavran-Greenberg, CB#7420, Chapel Hill, NC 27599, *fwright@bios.unc.edu*; Arpita Ghosh, The University of North Carolina at Chapel Hill; Fei Zou, The University of North Carolina at Chapel Hill

Key Words: genome scan, bias, association, genetic risk

It is widely recognized that genome-wide association studies suffer from inflation of the risk estimates for genetic variants (usually SNPs) identified as significant in the genome scan. To handle such significance bias, a number of investigators have proposed using likelihoods which condition on the declared significance of the outcome. We elaborate on an approximate conditional likelihood approach that can be applied using odds ratio estimates provided by standard statistical software, and consider competing estimators to the maximum of the conditional likelihood. We also discuss extensions to the situation where risk estimation is performed for multiple correlated phenotypes in the genome scan. The results have considerable importance for the proper design of follow-up studies and risk characterization.

On the Analysis of Copy-Number Variations in Genome-Wide Association Studies: A Translation of the Family-Based Association Test

✱Christoph Lange, Harvard School of Public Health, 02115, *clange@hsph.harvard.edu*

Key Words: Family-based association test, Copy Number Variants, Asthma

In this paper, we propose a generalization of family-based association tests (FBATs) to allow for the analysis of CNVs at a genome-wide level. We translate the popular FBAT-approach so that, instead of genotypes, raw intensity values that reflect copy number are used directly in the test statistic, thereby bypassing the need for a CNV genotyping algorithm. Both inherited and de novo CNVs can be analyzed without any prior knowledge about the type of

CNV, making it easily applicable to large scale association studies. All robustness properties of the genotype FBAT-approach are maintained and all previously developed FBAT extensions, including FBATs for time-to-onset, multivariate FBATs, and FBAT-testing strategies, can be directly transferred to the analysis of CNVs. Using simulation studies, we evaluate the power and the robustness of the new approach.

44 Memorial Session for Randy Sitter

Memorial, Section on Nonparametric Statistics, Section on Survey Research Methods, Section on Physical and Engineering Sciences, Section on Quality and Productivity, Committee on ASA Archives and Historical Materials
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Randy Sitter's Contributions to Resampling Methods in Sample Surveys

✱J. N. K. Rao, Carleton University, School of Mathematics and Statistics, Ottawa, ON K2G 4H8 Canada, *jrao34@rogers.com*

Randy Sitter has made fundamental contributions to resampling methods in sample surveys. In particular, he has developed ingenious methods for variance estimation under stratified multistage cluster sampling designs using balanced repeated replication and bootstrap sampling. In this talk, I will highlight some of Randy's major contributions to this important topic in survey sampling.

A Statistical Approach to Quantifying the Elastic Deformation of Nanomaterials

✱C.F. Jeff Wu, Georgia Institute of Technology, School of Industrial and Systems Engineering, Atlanta, GA 30332, *jeffwu@isye.gatech.edu*

Accurate estimation of elastic modulus of certain nanomaterials like nanobelt is important in many applications. A recently proposed approach was to estimate elastic modulus from a force-deflection model based on the continuous scan of a nanobelt using an Atomic Force Microscope tip at different contact forces. However, the nanobelt may have some initial bending and it may shift or deform during measurement leading to bias in the estimation. In this work we propose a new approach "Profile adjustment and parameter estimation (PAPE) to account for these various possible errors. It can automatically detect and remove the systematic errors and thus gives a more accurate estimate of the elastic modulus. The advantages of the approach are demonstrated through the application on several data sets. (Joint work with X. Deng, R. Joseph, W. Mai, Z. L. Wang.)

Using Empirical Likelihood Methods To Obtain Range Restricted

✱Jiahua Chen, University of British Columbia, Department of Statistics, Vancouver, BC V6T 1Z2 Canada, *jhchen@stat.ubc.ca*; J. N. K. Rao, Carleton University; Randy Sitter, Simon Fraser University

Design weights in surveys are often adjusted to accommodate auxiliary information and meet pre-specified range restrictions, typically via some ad hoc algorithmic adjustment to a generalized regression estimator. We present a simple solution to this problem using empirical likelihood methods or generalized regression. We first develop algorithms for computing empirical likelihood estimators and model-calibrated empirical likelihood estimators. The first algorithm solves the computational problem of the empirical likelihood method in general, both in survey and nonsurvey settings, and theoretically guarantees its convergence. The second exploits properties of the model calibration method and is particularly simple. The algorithms are adapted to handle benchmark constraints and prespecified range restrictions on the weight adjustments.

45 Training Community Collaborators To Understand and Effectively Use Health-Related Data ▲

Committee on Outreach Education, Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Training Community Collaborators To Understand and Effectively Use Health-Related Data

✱Francine Romero, Southwest Tribal Epidemiology Center, Albuquerque Area Indian Health Board, 5015 Prospect Avenue NE, Albuquerque, NM 87110, *fromero@aastec.org*; ✱Kapuaola Gellert, Na Pu'wuai Native Hawaiian Health Care System, Moloka'i, HI , *kapuaola@napuuwai.com*; ✱Walter T. Ambrosius, Wake Forest University School of Medicine, Department of Biostatistical Sciences, Winston-Salem, NC 27157-1063, *wambrosi@wfubmc.edu*

Key Words: community research, health data, teaching, native cultures, training, outreach

Many research projects and programs are conducted via collaborative partnerships between health agencies, academic institutions and community-based organizations. One widely used framework for these is Community Based Participatory Research (CPBR), which involves community members in "every facet of the research process." If community members are to have substantial input into research design, analysis, and interpretation of their data, they need to be able to understand and access the information in a manner useful for their purposes. However, lack of formal training in statistics makes it difficult for some of our collaborators to use the information gained from projects conducted in their communities. In this session, some solutions will be presented based on a variety of experiences of the panel members as they work in diverse communities and populations.

46 The Impact of AP Statistics on Our Profession: Past, Present, and Future ●▲

Mu Sigma Rho, Section on Statistical Education
Sunday, August 3, 4:00 p.m.–5:50 p.m.

The Impact of AP Statistics on Our Profession: Past, Present, and Future

✱Christine Franklin, The University of Georgia, Department of Statistics, 204 Statistics Building, Athens, GA 30602, *chris@stat.uga.edu*; ✱Richard Scheaffer, University of Florida, PO Box 118545, Gainesville, FL 32611, *rls907@bellsouth.net*; ✱Roxy Peck, California Polytechnic State University, Department of Statistics, San Luis Obispo, CA 93407, *rebecca@amstat.org*; ✱Katherine Tranbarger, Amherst College, Amhert Box 2239, Amhert,

MA 01002, *ktranbarger@amherst.edu*; ✱David Thiel, Clark County School District, 3950 S. Pecos-McLeod, Ste. 2-C, Las Vegas, NV 89121, *dthiel@interact.ccsd.net*

Key Words: AP Statistics, Impact on the profession, K-12 Statistics

The AP Statistics program celebrates the administration of the 12th AP Statistics exam in 2008. The program has grown from 7500 exams in 1997 to a projected 110,000 exams in 2208. What has been the impact of AP Statistics during this 12-year evolution? Where does the program stand today? What is the vision for the future of AP Statistics? The panel will discuss the past, present, and future of AP Statistics and it impact on educating young statisticians. This discussion will include why offer the AP Statistics program, the impact of AP Statistics on statistics at the Pre-K–12 level and on undergraduate and graduate program in statistics, the critical issues facing the AP Statistics program for the future success of statistics as a discipline, and how statisticians can become more involved to reach out and further the success of the AP Statistics program.

47 Clustering in Marketing Research: International Experience ●▲

Section on Statistics and Marketing, Social Statistics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Robust Model-Based Method for Cluster Analysis: Theory and Practice

✱Ewa M. Nowakowska, GfK Polonia, Smulikowskiego 4, Warsaw School of Economics, Warsaw, 00-389 Poland, *ewa.nowakowska@gfk.com*; Agnieszka Fronczyk, GfK Polonia; Krzysztof Z. Puszczak, GfK Polonia

Key Words: outliers, segmentation, ML-estimator

Outliers are observations substantially different from the rest of the data. Their undiscovered presence in the dataset is likely to result in misleading conclusions. The problem is vital for cluster analysis, where undetected contamination usually causes significant decrease in stability and reproducibility, leading to distortions in the segments' image. Based on the ML-approach proposed by Gallegos and Ritter (2005), we have implemented an effective algorithm, which allows for simultaneous detection of outliers and estimation of the structure in the remaining data. In contrast to the standard outlier detection procedure (applied prior to the segmentation itself), the method advantage is that it detects contamination in relation to the identified data structure. This presentation discusses the method applications in marketing research.

A New Approach for International Market Segmentation

✱Shon Magnan, GfK Custom Research North America, *shon.magnan@gfk.com*

Key Words: Marketing, Segmentation, Multi-National

Many multinational organizations are attempting to use market segmentation in the formulation of marketing strategies. International segmentation, however, is fraught with challenges. Should a bottom-up approach be used where each country is segmented separately from other countries? Or, should a top-down method be used where multiple countries are included in a single analysis? Neither approach typically will be optimal. A bottom-up approach ignores the global nature of business and results in multiple segmentation models that are difficult to implement. By focusing on a top-down approach, country nuances may be ignored. Another impediment to international segmentation is the issue of scale-use differences. This presentation introduces an approach for international segmentation that balances the need both for a global and local segmentation while accounting for scale-use.

Outlier Detection in Segmentation Studies

✱Norbert Wirth, Advanced Analytics Europe, MarketTools, 67-74 Saffron Hill, London, International EC1N 8QX United Kingdom, *Norbert.Wirth@markettools.com*

Key Words: Segmentation, Outlier, Outlier detection

Per definition, outliers are a part of the sample that originates from a different parent population than the rest of the sample. In practice outliers are typically seen as cases with biased or strongly atypical profiles. In the development of market segmentations such cases contaminate the entire process and reduce the quality of the solution. This paper discusses the impact of outliers on model and distance based cluster solutions, reflecting various potential data structures. Specific outlier constellations and their impact on the quality of results will be demonstrated and categorized, also considering interpretational consequences. Reflecting the drawbacks of common practices to handle outliers, a pre clustering outlier detection approach will be presented, suitable to minimize the negative impact on the resulting segments.

Customer Profiling in Mixed Models: Data Modeling with Cluster Analysis Results

✱Erin McClintic Tanenbaum, Claritas, A Nielsen Company, 1525 Wilson Blvd, Suite 1200, Arlington, VA 22209, *erin.tanenbaum@nielsen.com*

Key Words: cluster analysis, mixed models, marketing

In marketing, understanding the characteristics and preferences of customers through statistical approaches is an effective way to target customers and increase retention. Claritas employs demographic information coupled with consumer behaviors to create several customer profiling tools. The cluster results were employed by the analytic challenge winners at the most recent Direct Marketing Association (DMA) conference. This leaves the question: what is the proper way to use cluster analysis data in modeling? What are some common mistakes used when modeling with cluster analysis results? The presentation will be specific to marketing applications, but may be employed wherever mixed models are used.

Bayesian Networks in Segmentation for Marketing Research

✱Benoit Hubert, GfK, 10 rue Lionel Terray, Rueil-Malmaison, International 92508 France, *benoit.hubert@gfkcr.fr*

Key Words: Bayesian Networks, Clustering, Multiple Clustering, Segmentation, Causal Model

It has become recognized increasingly that Bayesian Networks (BN) are able to provide very good solutions to many actual market research problems. The objective of a BN is to encode the joint probability distribution of a set of random variables. Its graphical presentation is very useful and simple for exploration of relationships between phenomena. From expert specifications to automatic algorithms, the BN creation is now entering a quite mature phase. In this paper, we will focus on traditional algorithms that let people establish BN models from sets of data (Taboo, EQ, Augmented Naives Bayes, Markov blanket algorithms). We will describe a few ways to increase the power of BN to solve market research traditional problems especially in clustering and segmentation area. We will also present an example of a causal model that builds a link between BN and Structural Equation Modeling.

48 Recent Developments in Method and Application in High-Dimensional 'Omics' Data ●▲

Committee on Applied Statisticians
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Shrinkage Approach to Gene-Set Analysis: SAGA

✱Daniel C. Parks, GlaxoSmithKline R&D, 1250 S. Collegeville Rd, Collgeville, PA 19426, *daniel_c_park@gsk.com*; Xiwu Lin, GlaxoSmithKline R&D; Kwan Lee, GlaxoSmithKline

Key Words: Gene set analysis, Shrinkage approach

Research in gene set analysis of expression data using the gene ontology has drawn wide attention recently. The methods currently used, such as Functional Class Scoring (FCS, Pavlidis et al., 2004) and Gene Set Analysis (GSA, Efron and Tibshirani, 2006) adopt a univariate approach or semi-multivariate approach. While especially FCS can handle correlation within each gene set, they may not deal with the high correlation among gene sets that arise from many genes being annotated by several gene sets. This can lead to misinterpretation due to false-positive results. To correct for this, we propose a multivariate method entitled the Shrinkage Approach to Gene-set Analysis (SAGA), which can handle high dependence among gene sets because the shrinkage partial correlation cooperates with all variables. We compare SAGA to currently used method (GSA) using simulations and real data sets.

Estimation Equation Based Causality Analysis with Application to Microarray Time Series Data

✱Jianhua Hu, The University of Texas MD Anderson Cancer Center, Houston, TX , *JHu@mdanderson.org*; Feifang Hu, University of Virginia

Key Words: Chi-square approximation, Estimating equation, F-test, False positive rate, Granger causality, Time course data

Microarray time course data can be used to explore the interaction among genes. We are particularly interested in the causality network that is built by treating each gene's expression profile as a time series. A typical existing method establishes the Granger causality based on Wald-type of test, which relies on the homoscedastic normality assumption of the data distribution. However, this assumption can be seriously violated in real microarray experiments that may possibly result in false scientific conclusions. To overcome the drawback, we propose a method based on estimation equations which is robust to both heteroscedasticity and non-normality of the gene expression data. In fact, it only requires the residuals to be uncorrelated. We will use simulation studies and a real data example to demonstrate the applicability of the proposed method.

Non-Negative Matrix Factorization of Microarray Data

✱Katja S. Remlinger, GlaxoSmithKline R&D, PO Box 13398, Research Riangle Park, NC 27709-3398, *katja.s.remlinger@gsk.com*; S. Stanley Young, National Institute of Statistical Sciences; Paul Fogel, Consultant; Kejun (Jack) Liu, OmicSoft

Key Words: non-negative matrix factorization, microarray, clustering, multiple testing, variable reduction

Singular value decomposition is the basis of many linear modeling statistical methods. Often the data under analysis is non-negative, e.g. RNA expression levels, protein expression levels, metabolite expression levels, word counts in documents, gray scales in pictures, etc. Should the data analysis method take cognizance of the non-negative nature of the data? For non-negative matrices, non-negative matrix factorization, NMF, can replace SVD as the core computational method for data analysis. NMF is used to analyze data sets from the NCBI GEO database.

Segmentation of Treatment Effect via Analyzing Genetic Marker Profiles

✱Peter Hu, Merck & Co., Inc, 19454, *peter_hu@merck.com*

Key Words: hyper-responders, power, composite biomarker, personalized medicine, validate, response differentiation

Targeting the patient population for hyper- and hypo-responders, through either maximizing drug efficacy or alleviating negative reactions from a candidate drug, is an effective means to reach the ultimate goal of "personalized medicine." Two phase III clinical trials were utilized as an example to demonstrate the challenges of constructing composite biomarkers to achieve response differentiation. By carefully choosing appropriate algorithms, we demonstrated how response differentiation patterns can still be identified under circumstances when 1) there is limited sample size and 2) the power of individual markers differentiating responses is low. This talk will be focused on how to filter out noisy markers from consideration and how to avoid bias during selection of algorithms to build a conditional parametric regression tree for segmenting drug efficacy.

Analysis of Multiple Cancer GEO Microarray Data Sets

✱George Luta, Georgetown University, Washington, DC , *gl77@georgetown.edu*; Paul Fogel, Consultant; Kejun (Jack) Liu, OmicSoft; S. Stanley Young, National Institute of Statistical Sciences

Key Words: cancer mechanisms, microarray, meta analysis, multiple testing, confirmatory analysis

Over 7,000 microarray data sets are readily available from the NCBI GEO database, and many hundreds among them are aimed at unraveling important biological cancer mechanisms. Highly efficient software is now available for the analysis of microarray data, (e.g., Array Studio from OmicSoft). For these complex, n<<p experiments, there is a current need to study the statistical characteristics of both confirmatory and combined analyses of multiple datasets aimed at answering the same biological question(s). Our idea is to use the GEO database to identify complementary experiments and then perform both separate and combined analyses, the later with the potential advantage of increased statistical power.

49 Statistics in Ophthalmology Clinical Trials ●

Biopharmaceutical Section, WNAR, Biometrics Section
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Are Visual Acuity Measurements Independent of Test Distance? Statistical Analysis in Search for an Answer

✱Li Ming Dong, State University of New York at Stony Brook, Department of Preventive Medicine, Health Science Center-Level 3, Room 086, Stony Brook, NY 11794-8036, *LiMing.Dong@stonybrook.edu*

Key Words: Best-corrected visual acuity, test distances, ophthalmologic clinical trials

Best-corrected visual acuity (VA) is the most important outcome measure in ophthalmologic clinical trials. Commonly measured using the ETDRS charts, VA often had to be tested at multiple distances to accommodate the full range VA due to the size limit of the chart. VA scores from different distances were then converted to a common scale for analysis assuming that same visual angle results in same VA regardless of test distance. However, discrepancy in VA measurements from two distances was observed in four clinical trials. This discrepancy could lead to bias in estimating treatment effect. In this talk, I will first describe VA measurement and test protocol in ophthalmologic clinical trials. Then I will present statistical analyses conducted and issues rose when searching for reasons that might account for the observed discrepancy in VA measurements obtained from two test distances.

Statistical Analysis of Visual Field and Visual Field Loss Data in Ophthalmic Research

✱Fei Yu, University of California, Los Angeles, Los Angeles, CA 90095, *fyu@ucla.edu*; Anne L. Coleman, University of California, Los Angeles

Key Words: statistical analysis, visual field, glaucoma, Humphrey field analyzer

Visual field is an important aspect of visual function. The evaluation of visual field can provide essential information on the detection and progression of many eye diseases, such as glaucoma. Currently, automated computerized perimetry is the most widely-used method for visual field examination among eye care professionals, and is most commonly performed by using the Humphrey Field Analyzer because of its standardized output and in-depth statistical analysis of test results with its STATPAC software. Therefore, it is important to understand how the visual field is measured, the reliability measurements of visual field test (fixation loss, false positive errors and false negative errors), and the global summary of visual field test (mean deviation [MD], pattern standard deviation [PSD] and corrected pattern standard deviation [CPSD]).

Accounting for Correlation in Regression Analyses of Longitudinal Data from Paired Eyes

✱Gui-shuang Ying, University of Pennsylvania, 3535 Market Street, Suite 700, Philadelphia, PA 19104, *gsying@mail.med.upenn.edu*; Maureen Maguire, University of Pennsylvania

Key Words: multiple-level correlated data, generalized estimating equations, generalized linear mixed model, bivariate linear mixed model, Ophthalmology

In laboratory and clinical studies in ophthalmology, two eyes may be in the same or different treatment groups and followed longitudinally with repeated measurements. The analysis of such data is challenged by two-levels of correlation: one from repeated measurements over time from the same eye, and another from the paired eyes of the same subject. We discuss three analysis approaches including: generalized estimating equations (GEE), generalized linear mixed model (GLIMMIX), and bivariate linear mixed model. We compare their performance relative to the naïve approach that ignores correlation by Monte Carlo simulations under various sample sizes, degrees of correlation, and membership in the same or different treatment groups. Data from the Complications of Age-related Macular Degeneration Prevention Trial (CAPT), are used to illustrate the application of the methods.

Statistical Considerations in Ophthalmic Device Trials

✱Hollington T.C. Lu, U.S. Food and Drug Administration, 1350 piccard Dr. HFZ-550, Rockville,, MD 20850, *Hollington.lu@fda.hhs.gov*

Key Words: Ophthalmic clinical trials, Statistical analysis, medical device trial

The eye has a wide variety of structure in which physical and functional defects can have serious consequences impacting the ability to perform in everyday life. There are many ophthalmic device that have been developed in recent times for treating and managing eye diseases, including contact lens, phakic IOL, aphakic IOL, laser, ophthalmic viscoelastic device (OVD), nitrogen gas, silicon oil, and implantable vision enhance devices. For a biostatistician, the most important distinction between ophthalmologic data and data obtained in most other medical specialties is whether the information is collected on an eye-specific or subject-specific basis. The statistical issues and examples presented are the correlation of left and right eye, longitudinal nature of the ocular measurements, post-marketing study of longitudinal cell density and glaucoma diagnosis.

50 Statistical Methods in Public Health: Applications, Research, and Education ●

Section on Health Policy Statistics, Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Evaluating the Uncertainty in HIV/AIDS Infection Models

✱Huaiyu Ma, GE Global Research, One Research Circle, Niskayuna, NY 12309, *mah@research.ge.com*; Roger W. Hoerl, GE Global Research

The World Health Organization and many other NGOs utilize HIV infection models in order to make predictions and allocate funding. Examination of these models reveals that numerous assumptions tend to be made, such as model forms, model parameters, and distributions. While such assumptions are often necessary when modeling complex phenomena, we find that the uncertainty introduced by these assumptions has not been adequately quantified, and may be seriously underestimated. We demonstrate one approach to evaluating and quantifying the uncertainty associated with these models. It is our belief that a more accurate understanding of the magnitude of the uncertainty will result in better decisions being made using these models.

Recent Background and Influences on the Direction of Research in Public Health Biosurveillance

✱Henry Rolka, Centers for Disease Control and Prevention, *hrolka@cdc.gov*

Key Words: biosurveillance, situational awareness, public health

Biosurveillance for Public Health has been a focus of recent legislative and policy initiatives such as the Pandemic and All Hazards Preparedness Act and Homeland Security Presidential Directives. The initiatives produce increased technical requirements for data analysis, information science and communicating biosurveillance results involving the characterization of uncertainty. In addition to traditional analytic epidemiological studies, data and information processing is used to establish a “common operating picture” and “situational awareness” for public health on an ongoing bases and especially to be used in responding to emergencies. These evolving public health operational requirements augment the complexity of practical biosurveillance. This talk will provide a brief background and history with some examples of challenging components and areas of potential research.

Statistical Considerations in Large-Scale Screening Programs: Impacts on the Public

✱Sonya Heltshe, Rocky Mountain Denver Poison Center, *sonyaheltshe@gmail.com*; Karen Kafadar, University of Colorado at Denver and Health Sciences Center; Philip Prorok, National Cancer Institute

We describe some of the challenges involved in planning, conducting, and analyzing data from, randomized screening trials that distinguishes them from therapy clinical trials, including issues of lead time, overdiagnosis, and length bias. The present PLCO screening trial, designed to evaluate the effect of screening for prostate, lung, colorectal, and ovarian cancers, will illustrate the issues. The impact of screening on the public, including financial, sociological, and health aspects, will be discussed.

Assessing the Effects of Air Pollution on Human Health

✱Rebecca Klemm, Klemm Analysis, Inc., 1339 Green Court, NW, 4th Floor, Washington, DC 20005, *rklemm@klemmanalysis.com*; Donna Stroup, Data for Solutions, Inc.

We examine methodology for developing and validating models to estimate the effect of air pollution (particulate matter, PM) on health. Changes in PM depend on measurement techniques, modeling, and analysis. We use NCHS mortality data on respiratory and cardiovascular disease, and develop and innovative method for assigning distance between place of death and air quality monitoring source. Models to predict mortality use Air quality data from EPRI and weather data from NOAA. We show the value of information from multiple causes of deaths due to respiratory conditions, unintentional injuries, septicemia, and Alzheimer's disease emphasizes the importance of statistics in guiding the proper use and interpretation of model effects on outcomes.

51 Semiparametric Bayesian Methods

Section on Bayesian Statistical Science, Section on Nonparametric Statistics

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Toward Fully Nonparametric Bayesian Regression

✱Siddhartha Chib, Washington University in St. Louis, CB 1133, 1 Brookings Dr, St. Louis, MO 63130, *chib@wustl.edu*; Edward Greenberg, Washington University in St. Louis

Key Words: additive regression, cubic spline, Dirichlet process mixture, marginal likelihood, Markov chain Monte Carlo

We present a nonparametric Bayesian regression approach in which covariate effects are modeled additively by cubic splines and the error distribution by a Dirichlet process prior. This represents a generalization of existing models that contain only one of these features. One notable aspect of our approach is the adaptation of a cubic spline basis that was suggested in the literature dealing with interpolation methods but has been little used in statistics. We discuss a systematic approach for specifying proper prior beliefs on the coefficients of the basis functions and show how the model can be estimated efficiently by MCMC methods. Comparisons with (partial) nonparametric models are included. All the ideas are illustrated with simulated and real data.

Bayesian Estimation of ROC Curve Based on Ranks

✱Subhahsis Ghoshal, North Carolina State University, Department of Statistics, 2501 Founders Drive, Raleigh, NC 27695, *ghoshal@stat.ncsu.edu*; Jiezhun Gu, Duke Cliinical Research Institute

Key Words: ROC curve, rank likelihood, posterior consistency, Gibbs sampling, gold standard

The ROC curve, measuring disparity of two distributions, is extremely valuable in diagnostic testing and radiology. Under the popular binormal model, the two sets of observations follow two different distributions which are normal up to some unknown monotone increasing transformation H. The ROC curve does not depend on the nuisance H. We propose a rank based likelihood completely free of H, and adopt a Bayesian approach. We describe Gibbs sampling strategy and show posterior consistency. When the labels of the observations are not observed (commonly known as the case without the gold standard for the truth), we either put a prior on H based on Bernstein polynomials, or use latent variables for the lost labels. Simulations show that both methods give more reliable results than the commonly used method based on a surrogate marker.

Modeling Heteroscedasticity in the Single-Index Model with the Dirichlet Process

✱George Karabatsos, University of Illinois-Chicago, College of Education, 1040 W. Harrison St. (MC 147), Chciago, IL 60607-7133, *georgek@uic.edu*

Key Words: Bayesian nonparametrics, Ridge regression, dimension reduction, Gibbs sampling, adaptive Metropolis-Hastings

In the single-index model, a nonparametric regression model which has seen many applications, it is assumed that each observed continuous response has mean equal to an unknown (link) function of a linear combination of p regression coefficients and covariates, and assumes homoscedasticity. The model avoids the curse of dimensionality, and automatically handles interactions among the p covariates. A novel Bayesian heteroscedastic single-index model is introduced, where the link function is modeled by splines that are regularized by a ridge prior with parameter assigned a hyperprior, and where the distribution of the error variances (over observations) is modeled by a nonparametric, Dirichlet Process prior. The heteroscedastic model is illustrated through the analysis of real and simulated data, and it demonstrated superior predictive accuracy, compared to the homoscedastic model.

Modeling with Double Dirichlet Process Mixtures

✱Sanjib Basu, Northern Illinois University, DU 366, Division of Statistics, De Kalb, IL 60115, *basu@niu.edu*; Siddhartha Chib, Washington University in St. Louis

Key Words: Dirichlet process mixtures, importance sampling, model comparison, Markov chain sampling

In this work, we consider variants of the Dirichlet process mixture model, that we call the double and multiple DPM models. These models have several attractive features. For one they are easy to interpret. Fitting of these models is straightforward by Markov Chain sampling. The models can also be estimated by a novel importance sampling scheme. We show how this scheme can be used for model comparison, thus extending the work of Basu and Chib (2003) on single DPM models. We illustrate the implementation, applicability and performance of these models with simulated and real data on binary and clustered outcomes.

Parametric and Semiparametric Hypotheses in the Linear Model

✱Steven N. MacEachern, The Ohio State University, Department of Statistics, 1958 Neil Ave., Columbus, OH 43210-1247, *snm@stat.osu.edu*; Subharup Guha, University of Missouri-Columbia

Key Words: Bayes factor, Dirichlet process, Kullback-Leibler divergence, nonparametric Bayes

The independent additive errors linear model consists of a mean structure and a separate error structure. The error structure may be parametric or semiparametric. Under alternatives for the mean structure, the best fitting additive errors model has an error distribution which can be represented as the convolution of the actual error distribution and the marginal distribution of a misspecification term. The misspecification term follows from the covariates' distribution. Conditions are developed to distinguish when the semiparametric model yields sharper inference than the parametric model and vice-versa. The results explain a paradoxical finding in semiparametric Bayesian modeling, where the posterior distribution under a semiparametric model is found to be more concentrated than under a corresponding parametric model. The paradox is illustrated with a standard regression data set.

52 Current Statistical Topics Involving Individual Income Tax Statistics

Section on Government Statistics, Section on Survey Research Methods, Section on Quality and Productivity, Social Statistics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Statistics from Individual Income Tax Returns: Populations, Samples, and Processing of Individual Income Tax Returns at Statistics of Income

✱Michael E. Weber, Statistics of Income, IRS, P.O. Box 2608, Washington, DC 20013-2608, *michael.e.weber@irs.gov*; David P. Paris, Statistics of Income, IRS; Peter J. Sailer, Consultant

Key Words: Tax Returns, Administrative Records

Statistics from Individual Income Tax Returns have been produced by the Statistics of Income Division since 1917. This paper discusses the statistics generated from the yearly filing of Individual Income tax returns from the early 1960s to the present. It traces changes made to the yearly sampling plan from a single cross section used for national estimates to the current configuration which includes a cross section for national estimates, a Continuous Work History Sample panel, multiple stratified high income cohort panels, and a an expanded cross section sample for state level estimates, as well a nonfiler sample. The paper will also explain how these samples are processed from raw administrative data into perfected data as well provide definitions of SOI terminology and a description of the various products derived from these samples.

SOI/IRS Sales of Capital Assets Sample Redesign for Tax Year 2007

Yan K. Liu, Statistics of Income, IRS; ✱Michael Strudler, Statistics of Income, IRS, *Michael.i.strudler@irs.gov*; Jana Scali, Statistics of Income, IRS; Janette Wilson, Statistics of Income, IRS

Key Words: Neyman allocation, stratified sample

The Statistics of Income (SOI) of the IRS developed a stratified sample of individual returns to study the form 1040 Sales of Capital Assets (SOCA) in tax year 1999. It was a cross-sectional sample from the population of all individual returns of tax year 1999. Because of high processing cost of SOCA returns, there has been no other SOCA cross-sectional sample ever since. However, the 1999 SOCA cross-sectional sample is outdated as there have been many economic changes that impact capital gains. Therefore, it is decided to start a new SOCA cross-sectional sample for tax year 2007. This paper discusses how the new sample is designed, including determining stratum boundaries, allocating sample sizes to strata, and balancing the cost and precision. The Neyman allocation is used and the information on cost estimate and variance estimate is obtained from the related SOI's data sources.

Attrition in the Tax Years 1999–2005 Individual Income Tax Return Panel

✱Victoria Bryant, Statistics of Income, IRS, P.O. Box 2608, Washington, DC 20013, *Victoria.L.Bryant@irs.gov*

Key Words: Income Tax, Panel, Attrition

Policy research is increasingly being done on panel data; attrition can undermine validity and misrepresent results of many policy analyses. Using the Individual Income Tax Return Panel for Tax Years 1999-2005, this paper will examine panel attrition. It tests the randomness of attrition and evaluates the implications of such results. It assesses the observed rate to determine the predictability over time. Finally, it will present several tabular representation problems hindering analysis.

Methodological Limitations in Producing Subnational Tabulations of Unincorporated Business Activity That Partnerships and Sole Proprietorships Report on Returns

✱Joseph Koshansky, Internal Revenue Service, *joseph.koshansky@irs.gov*

Key Words: Statistics of Income, Unincorporated business activity, Subnational estimates

The Statistics of Income (SOI) Division generally compiles statistics based on stratified probability samples, using such classes as size of income, presence or absence of a specific form or schedule, and business activity. In addition, it evaluates these estimates by comparing them with information it collects from extracts of the population of filers. This paper addresses the methodological limitations associated with producing state and county-level tabulations of unincorporated business activity from the information on a Form 1065 or a Form 1040, Schedule C. It presents trends and comparisons among states, counties and business activities across five tax years—2001 to 2005. Discussion includes the methodology for assigning entities to states and counties, as well as comparisons at the national level between this data and SOI estimates drawn from samples.

53 Recent Methodological Developments in Analysis of Large Data Sets

Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, Section on Statistical Computing

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Model-Free Variable Selection via Adaptive LASSO

✱Liqiang Ni, University of Central Florida, Department of Statistics, Univ. of Central Florida, Orlando, FL 32816-2370, *lni@mail.ucf.edu*

Key Words: canonical correlation, lasso, sliced inverse regression, B-splines, sufficient dimension reduction

Recently there has been considerable interest in variable selection via regularized methods mostly developed for single-index or semiparametric models. We propose a penalized objective function using transformations of the response as the multivariate pseudo-response and utilizing an adaptive lasso as the operator. Under mild assumptions on the marginal distribution of the predictors, the proposed approach selects variables consistently without restrictive model assumptions. Simulations confirm these findings.

Central Mean Subspace in Time Series

✱Jin-Hong Park, College of Charleston, Robert Scott Small Building / Room 351, 66 George St., Charleston, SC 29424, *parkj@cofc.edu*

Key Words: Time series central mean subspace, Ordinary least square, Kernel method, Schwarz (Baysian) information criterion, Residual information criteria

We propose a method of dimension reduction in the mean function of a time series, which does not require specification of a model but seeks to find a p*d matrix with the smallest possible number d (< p) such that the conditional mean is same as the expected value of current given past linear combinations, resulting in no loss of information about the conditional mean of the series given its past p values. We propose notions of time series central mean subspace and minimum mean dimension reduction subspace, and estimate the latter using a method based on residual sum of squares, when p and d are known. In order to estimate the correct dimension, we introduce two information criteria. We examine the performance of all the estimators extensively through simulation and real data analysis. As for estimation of the unknown lag, we use a graphical approach for visual appeal and simplicity.

On Parametric and Semiparametric Approaches to Differential Gene Expression Detection and Sample Size Calculation for Microarray Study

✱Baolin Wu, The University of Minnesota, A460 Mayo Building, MMC 303, Minneapolis, MN 55455, *baolin@biostat.umn.edu*; Dan Nettleton, Iowa State University

Key Words: microarray, multiple testing, FDR, experiment design

We study parametric and semiparametric approaches for large-scale simultaneous significance testing, which are also useful for future experiment design. Efficient computational algorithms will be proposed and compared to existing approaches. Simulations and applications will be conducted to illustrate the competitive performance of the proposed methods.

A Nonparametric Significance Test of No Trend for Functional Data

Jeongyoun Ahn, The University of Georgia; ✱Cheolwoo Park, The University of Georgia, Department of Statistics, University of Georgia, Athens, GA 30622, *cpark@stat.uga.edu*; Woncheol Jang, The University of Georgia

Key Words: False discovery rate, Functional data, Multiple testing, Principle component, Trend detection, Wavelet

We consider multiple testing of no trend in functional data. We test the null hypotheses that each function has no trend based on the Adaptive Neyman test and wavelet thresholding test introduced by Fan. Our methodology can be applied for nonstationary time series using principle component analysis and controls the false discovery rate to adjust multiplicity. We further extend the idea to functional analysis of variance and goodness-of-fit testing. The methods are illustrated with a simulation study and period change analysis of pulsating variable stars.

Sparse Distance-Weighted Discrimination

✱Lingsong Zhang, Harvard University, 655 Huntington Ave., Building 2, 4th Floor, Boston, MA 02115, *zhang@hsph.harvard.edu*; Xihong Lin, Harvard University

Key Words: Distance Weighted Discrimination, High Dimension Low Sample Size, proteomics, Sparsity, L_1 norm

In the High Dimension Low Sample Size situation, Marron et al (2002) proposed a new classification method, Distance Weighted Discrimination (DWD), which is similar to Support Vector Machine (SVM) when the number of objects is larger than the number of features, but perform better than SVM on high dimension low sample situation. However, in the high dimensional case, the noise in the data may still dominate in finding the separating hyperplane. In this paper, we proposed a Sparse DWD (SDWD) method, which incorporates the variable selection along with the classification. Theoretical properties are explored under some special conditions. Applications to proteomics and genetic pathways are used to illustrate the SDWD method.

54 Multivariate Outlier Detection: Applications and Methodology

Section on Statistical Computing, Section on Physical and Engineering Sciences

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Multivariate Outlier Detection by Partial Mixture Fits

✱David W. Scott, Rice University, Dept of Statistics, MS-138, POB 1892, Houston, TX 77005, *scottdw@rice.edu*

A new method for detecting multivariate outliers by fitting partial mixtures is proposed. Computational and performance issues are discussed, and the method is illustrated on real data.

Statistical Methods for Detecting Computer Attacks from Streaming Internet Data

✱Ginger Davis, University of Virginia, P.O. Box 400747, 151 Engineer's Way, Charlottesville, VA 22904, *gingerdavis@virginia.edu*; Karen Kafadar, University of Colorado at Denver and Health Sciences Center; David Marchette, Naval Surface Warfare Center

Successful strategies for network security depend upon collating massive data features into usable variables, high-performance and computationally efficient methods for anomaly detection, and ability to predict typical behavior so atypical and abnormal events can be better identified. We define data variables (packets, sessions, activities) from network data and how they can be used in models for user profiles, workload management, application verification, etc. New techniques to process these data will be discussed, including visualization of data sets, leading to methods for clustering images. We provide examples of these methods on motivating internet packet data streams.

High-Energy Physics: Massive Data Analysis in Action

✱Karen Kafadar, University of Colorado at Denver and Health Sciences Center, PO Box 173364, CB 170, Denver, CO 80217-3364, *karen.kafadar@cudenver.edu*; Robert Jacobsen, University of California, Berkeley

Experiments in high-energy physics provide terabytes of data, from which critical information about the state of matter, governed by the theory outlined in the "Standard Model," must be extracted. Opportunities abound for increased efficiencies in approaches to the data, from the design of experiments, to the collection of data, and finally to analysis and inference. Due to the massive amounts of data, from various sources (different experiments from different collaborations, experiment-based simulations, etc.), new ways of analyzing the data to answer questions of interest are devised. We describe the framework for these experiments and illustrate a new method for comparing multivariate (non-Gaussian) distributions.

Multivariate Analysis in Planetary Science: Understanding the Atmospheres of Jupiter and Saturn

✱Amy Braverman, Jet Propulsion Laboratory, Mail Stop 306-463, 4800 Oak Grove Drive, Pasadena, CA 91109-8099, *amy.braverman@jpl.nasa.gov*; Irina A. Kukuyeva, University of California, Los Angeles; Padma Yanamandra-Fisher, Jet Propulsion Laboratory, California Institute of Technology; Jan Deleeuw, University of California, Los Angeles; Amy Simon-Miller, NASA Goddard Space Flight Center

Very little is understood about the structure and dynamics of the atmospheres of the solar systems two largest gas giant planets: Jupiter and Saturn. In recent years a wealth of data has been collected from both satellite and ground-based instruments that promises to begin to shed light on these mysterious planetary bodies. Data sources include spectra obtained from the Hubble Space Telescope (HST), NASA Infrared Telescope Facility (IRTF) on Mauna Kea, and the Cassini spacecraft which is currently producing massive data sets as it continues to orbit Saturn. In this talk we will review some of the scientific questions of interest, and how multivariate analysis of remote sensing spectra may be used to help answer them.

Monitoring the Golden Gate Bridge Using Wireless Sensor Networks

✱Guilherme V. Rocha, University of California, Berkeley, 367 Evans Hall, Berkeley, CA 94720, *gvrocha@stat.Berkeley.EDU*

A 51-node Wireless Sensor Network have been installed on the Golden Gate Bridge for Structural Health Monitoring. There is a mismatch between the rate at which the data are collected at each node (~4 Kbps, Kilo-bytes/second) and the rate at which they can be transmitted (~0.5 Kbps). The ultimate goal is to develop a data reduction scheme so the WSN can perform real time monitoring of the dynamic properties of the bridge. For the temperature data (~0.80 Kbps/sensor), lossless run length coding achieves a significant reduction in the data rate (~0.04 Kbps/sensor). For the acceleration data, our strategy is to construct a parametric model for the bridge and continuously adjust it as data become available. We report the results of such strategies for both simulated and actual data. Joint work with David Culler, James Demmel, Gregory Fenves, Sukum Kim, Shamim Pakzad, and Bin Yu.

55 Clickers in the Statistics Classroom ●▲

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Clickers in the Statistics Classroom

✱Mark Berenson, Montclair State University, Dept. of Management & Information Systems, Montclair, NJ 07043, *berensonm@mail.montclair.edu*; ✱Deborah V. Dawson, The University of Iowa, Dows Institute for Dental Research, Room 439E Dental Science Building North, Iowa City, IA 52242-1010, *deborah-dawson@uiowa.edu*; ✱Herle McGowan, The University of Michigan, 439 West Hall, 1085 South University, Ann Arbor, MI 48109-1107 United States, *hmcgowan@umich.edu*; ✱Robert Terry, University of Oklahoma, Department of Psychology, 808 DAHT, Norman, OK 73019, *rterry@ou.edu*

Key Words: statistical education, pedagogy, technology, clickers

Professors from a diverse group of universities and disciplines will discuss the use of clickers in their applied statistics courses and what their experiences, both positive and not-so-positive, with clickers have been. Discussion will cover many of the choices in implementing a clicker system, including what systems are available, the nature and number of clicker questions and issues in grading. By the end of this session, attendees should have the opportunity to experience a clicker system and have a better idea of the benefits and limitations of this technology for teaching statistics.

56 Improving Survey Quality Reinterviews, Paradata, and Comprehension ●

Section on Survey Research Methods, Section on Government Statistics, Section on Quality and Productivity

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Reinterview Revisited

✱Dhirendra N. Ghosh, Synectics for Management Decisions Inc., 1901 N. Moore St., Suite 900, Arlington, VA 22209, *dhiren@smdi.com*

Key Words: Gross discrepancy rate, Response variance, Response bias, Index of inconsistancy (IOI)

The reinterview method is the usual technique for estimating the response variance and the response bias. It is assumed that the respondent does not remember or is not influenced by the original response. The gross discrepancy rate (GDR) is the ratio of the sum of the two types of discrepancies to the total number of respondents in a yes/no type of question. The response variance is half of the GDR. Response bias is also estimated from the ratio of the net difference between the two discrepancies and the total. However this assumption that the respondent neither remembers nor influenced by the original response is not realistic. A measure of the memory effect is derived in this paper. The GDR, the sample response variance and the response bias are rectified based on this measure of the memory effect.

Sample Designs for Evaluating the Quality of Data Collection

✱Cherie Alf, Center for Survey Statistics and Methodology, Iowa State University, Department of Statistics, Iowa State University, Ames, IA 50011-1210, *cjalf@iastate.edu*; Sarah Nusser, Center for Survey Statistics and Methodology, Iowa State University; Veronica Lessard, U.S. Department of Agriculture

Key Words: Data quality, In-stream sampling

A vital component in any survey is the quality assessment of data collection. Although it is standard practice for a subset of data collectors' cases to be monitored, sample designs are not always used to select cases for review. To be most effective, the quality review sample is chosen as the data collection is taking place, which calls for in-stream sampling methods. Several considerations go into designing an in-stream quality review sample, including which sample units to target, what information is available for targeting the selection process (e.g., sample frame information, survey responses, paradata from the interview process), temporal aspects of the sampling rate, and how to design the sample so that the data collector cannot predict when review will occur. We will consider this problem generally and provide an example from the National Resources Inventory.

Simulation Modeling of Field Operations in Surveys

✱Bor-Chung Chen, Federal Railroad Administration, 1200 New Jersey Avenue, SE, Washington, DC 20590, *Bor-Chung.Chen@dot.gov*

Key Words: Discrete Event Simulation, Field Operations, Survey Methods, Survey Costs, Response Rates

Simulation modeling has become the most commonly used tool for performance evaluation of stochastic dynamic systems in science and engineering. The field operations of surveys can be one of these systems. At Census Bureau, the Simulation Modeling of Field Operations in Surveys (SIMFOPS) was developed to estimate the cost, response rate, and timing of new or continuing surveys. In this paper, we describe the simulation of simplified field operations for NHIS (National Health Interview Survey). We apply the simulation methodology to the field operations. We use the 2004 NHIS CHI (Contact History Instrument) data for the study. From this study, we have shown that SIMFOPS can be used for optimizing the field operations by setting the controllable parameters before a decision is made and implemented. The cost savings might be enormous and would not be at the expense of the response rate.

Using Paradata to Actively Manage Data Collection

✱François Laflamme, Statistics Canada, Tunneys Pasture, Jean Talon Blg, 6e C-2, Ottawa, ON K1A 0T6 Canada, *francois.laflamme@statcan.ca*; Mike Maydan, Statistics Canada; Andrew Miller, Statistics Canada

Key Words: Paradata, call history, active management, responsive design, attempts

With the increasing complexity of survey design, survey managers need better tools to assess the data collection survey process. The development of data collection computer-assisted methods has opened the door to a wide scope of process data ("paradata") that can be used to evaluate the data collection survey process status in a timely manner. Active Management at Statistics Canada is a set of plans and tools developed to manage survey data collection while it is in progress. This includes monitoring progress, timely analysis of these indicators, identifying problems, implementing and communicating corrective actions, and evaluating success. The paper begins with an introduction to the key components of the Active Management and discusses a series of practical example currently used to monitor issues and data quality indicators at Statistics Canada.

A Microlevel Latent Class Model for Measurement Error in the Consumer Expenditure Interview Survey

✱Brian J. Meekins, Bureau of Labor Statistics, 2 Massachusetts Ave, NE, Washington, DC 20212, *meekins.brian@bls.gov*; Clyde Tucker, Bureau of Labor Statistics; Paul P. Biemer, RTI International/UNC-CH

Key Words: Latent Class Analysis, Measurement Error, Expenditure Reports

Previous research by Tucker et al. (2005) and Tucker et al. (2006) attempts to identify a latent construct that predicts the amount of measurement error in expenditure reports on the Consumer Expenditure Survey (CEIS). While this work was successful in identifying a construct that predicts measurement error in expenditure reports, it is more sensitive to falsely negative reports of the entire purchases than it is to the underreporting of the amount of expenditure for those purchases. Current research focuses more deeply on the underreporting of expenditure amounts for a number of different commodities. Together, with previously explored indicators such as, the number of contacts, missing on the income question, the length of the interview, and the use of records, we examine a new indicator measuring the number of missing sections of the interview, resulting in a new latent construct.

How To Prioritize Recontacts

✱Dan Hedlin, Statistics Sweden, Box 24300, Stockholm, SE-10451 Sweden, *dan.hedlin@scb.se*

Key Words: Selective editing, score function, survey, metric

Errors in data haunt practitioners. We have to prioritize observations that are most cost-effective to recontact and correct. One approach is to predict the true value of one observation using available information and take the norm of the difference between the predicted and reported value. This is a number referred to as an item score. Usually we want to edit and verify all observations on the same unit rather than each item separately. We discuss ways of forming a unit score out of a generic set of p item scores. A unit score function that unifies item scores often used in statistical editing is presented. Based on the unit score several subsample designs are available (e.g., pps or cutoff sampling above a threshold). We discuss ways of selecting this threshold value. The problem of prioritizing manual statistical editing of business survey data is our motivating example.

57 Sample Redesign

Section on Survey Research Methods, Social Statistics Section, Section on Government Statistics

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Sample Refreshment, Optimal Allocation, and Other Issues in the Redesign of Panel Surveys

✱Paul P. Biemer, RTI International/UNC-CH, 1917 Eagle Creek Ct., Raleigh, NC 27606, *ppb@rti.org*; Margaret Byron, Rho, Inc.; Sara Wheeless, RTI International; Celia Eicheldinger, RTI International

Key Words: longitudinal design, power analysis, NSCAW, child abuse or neglect, design effect, composite size measure

This paper summarizes research that was recently conducted for redesigning the NSCAW, a large, national panel survey. In 1998, a two-stage sample was selected consisting of approximately 5400 individuals in 92 PSUs. The sample continued unchanged for 8 years until it ended in 2007. In 2008, a new sample of 5700 individuals will be drawn to restart the survey. A number of issues were addressed for the new sample. Three options were considered for selecting new PSUs: (a) carry forward all 92 previously selected PSUs; (b)

replace all PSUs with an independently selected sample; and (c) select a high probability of overlap PSU sample. Other issues investigated include: optimal sample size, optimal sample allocation, frame changes, and domain oversampling while minimizing the design effects. This paper considers these issues and others and provides our recommendations for the redesign.

American Community Survey (ACS) Sample Design Issues and Challenges

✱Steven Hefter, U.S. Census Bureau, Washington, DC 20233, *steven.p.hefter@census.gov*

Key Words: ACS, Sampling, Census Long Form

The ACS collects data historically collected by the Decennial Census Long Form. The goal of the ACS is to publish data for small geographic areas cumulating sample over five-year periods with comparable quality to the Census 2000 Long Form estimates. Unlike the Long Form which had an overall sampling rate of approximately 1-in-6 and 100 percent follow-up of nonrespondents, the ACS selects a fixed annual sample of approximately 3,000,000 addresses and samples non-respondents with an overall rate of approximately 1-in-3. These two factors combined with growth and a decline in cooperation rates have led to concerns about the ACS meeting it's stated goal for small areas. This paper looks at the changes in the ACS sample size over time, the distribution of the ACS sampling frame relative to Census 2000 Long Form, and the implications of the fixed sample size on reliability.

Redesigning the American Community Survey (ACS): Computer-Assisted Personal Interview Sample

✱Edward C. Castro Jr., U.S. Census Bureau, Washington, DC 20233, *edward.c.castro@census.gov*; Steven Hefter, U.S. Census Bureau

Key Words: ACS, Sampling, Nonresponse, Personal Interview

The ACS collects data in three phases: mailout/mailback, telephone interview, and Computer Assisted Personal Interview (CAPI). During the CAPI phase, a sample of mail/telephone non-respondents and addresses deemed to be unmailable is selected for a personal visit. The sampling rate for this process is determined by the CAPI sampling stratum for the tract in which the address resides, which is assigned based upon the mail/telephone cooperation rate of the tract. The initial assignment was made prior to 2005. Based upon recent ACS data, cooperation rates were re-calculated for all tracts, which were then assigned to a new CAPI sampling stratum. This paper discusses the data that led us to revisit the CAPI sampling stratum assignments, and it provides the methodology used to calculate the tract level mail/telephone cooperation rates used to assign each area to a CAPI sampling stratum.

Sample Allocation and Relative Standard Error Calculation for National Hospital Discharge Survey Redesign

✱Iris Shimizu, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *ishimizu@cdc.gov*; Rong Cai, National Center for Health Statistics

Key Words: NHDS redesign, sample allocation, weighted intra-class correlation, relative standard error

The National Hospital Discharge Survey (NHDS) is being redesigned for 2010. The new design will use a two stage sample of hospital discharges in which a stratified sample of hospitals will be selected at the first stage. In this paper the hospital sample is allocated to strata by using Neyman and nonlinear programming methods. This research uses a sampling frame compiled from a 2006 list of hospitals which satisfy criteria for inclusion in the universe for the redesigned survey. For each stratum, weighted intraclass correlations are calculated for selected variables from 2005 NHDS data. The different allocations are evaluated by producing and comparing relative standard errors (RSEs, aka coefficient of variation or CVs) for those variables. The RSEs are approximated by using the relationship between variances and intra-class correlations.

Assessing Oversampling of Older Persons in a National Health Survey

✱Karen E. Davis, National Center for Health Statistics, 3311 Toledo Road, Rm 3213, Hyattsville, MD 20782, *kedavis@cdc.gov*; Chris Moriarity, National Center for Health Statistics

Key Words: Sample survey, Oversampling, NHIS

The National Health Interview Survey (NHIS) is one of the major data collection programs of the National Center for Health Statistics. The sample design for the NHIS traditionally has undergone a redesign about every 10 years to address new and continuing data needs at the national and subnational levels, and for minority and economic subdomains of the population. The ability to produce reliable annual estimates for older persons of age 65+ years by race and ethnicity is a major design objective. In 2002, research was conducted for the 2006–2014 NHIS redesign to assess options for oversampling older minority persons. Beginning with the 2006 NHIS, minority persons of age 65+ years have an increased probability of selection as sample adults. This paper describes research that has been conducted using the 2006 NHIS to assess the selection and the sample yield for older minority persons.

Implementation of Controlled Selection in the National Compensation Survey Redesign

✱Lawrence R. Ernst, Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 1950, Washington, DC 20212, *ernst.lawrence@bls.gov*; Christopher J. Guciardo, Bureau of Labor Statistics; Yoel Izsak, National Agricultural Statistics Service; Jonathan J. Lisic, Bureau of Labor Statistics ; Chester H. Ponikowski, Bureau of Labor Statistics

Key Words: allocation, sampling panels, rounding, sampling cells, area PSUs, industry sampling strata

The paper "Update of the Redesign of the National Compensation Survey" (Izsak et al. 2005) included a proposal to allocate the number of sample establishments among cells using a controlled selection procedure, where cells are area PSUs x industry sampling strata x sampling panels. Since then the procedure has been implemented but with a number of modifications not discussed in Izsak et al. (2005). These modifications and possible future changes are discussed in this new paper. They include: weighting changes necessitated by the use of controlled selection, complications caused by rounding issues and how they were overcome, complexities caused by the need to allocate over five sampling panels, and use of a real-valued minimum allocation for each sampling cell in the controlled selection process in order to avoid very large sample weights and accompanying increases in variances.

58 Models for Longitudinal Data ●

Section on Statistics in Epidemiology, Biopharmaceutical Section, WNAR, Section on Survey Research Methods, Biometrics Section
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Testing for Qualitative Interaction of Multiple Sources of Informative Dropout in Longitudinal Data

✱Sara B. Crawford, Valparaiso University, Valparaiso, IN 46383, *Sara.Crawford@valpo.edu*; John J. Hanfelt, Emory University

Key Words: Qualitative interaction, Informative dropout, Simultaneous confidence intervals, Clustered data

Longitudinal studies are subject to patient dropout, which may lead to biased estimates of the rate of change if the dropout process is informative. Multiple patterns of informative dropout are plausible, in that different causes of dropout might contribute to different patterns. These multiple patterns can be dichotomized into two groups: those exhibiting quantitative interaction and qualitative interaction. We explore a test for qualitative interaction based on simultaneous confidence intervals. The test accommodates the realistic situation where reasons for dropout are not fully understood, or perhaps are even entirely unknown. It also allows for an additional level of clustering among participating subjects, as might be found in a family study. We apply these methods to a longitudinal study exploring rates of change in cognitive functioning for a group of Alzheimer's patients.

Overdispersed Continuous-Time Models with Epidemiological Applications

✱Carles Breto, Universidad Carlos III de Madrid, C/ Madrid, 126, Statistics Department, Getafe, International 28903 Spain, *cbreto@est-econ.uc3m.es*; Edward L. Ionides, The University of Michigan; Aaron A. King, The University of Michigan

Key Words: overdispersion, continuous time, compartmental models, Markov processes, cholera, counting processes

Compartmental models are common in population modeling in the biological sciences but are also used in telecommunications, chemical kinetics and finance. In continuous time, standard compartmental models are based on 'Poissonian' processes which inherit the equidispersion property that the infinitesimal variance equals the infinitesimal mean. Recently developed techniques provide general methods for fitting such models to time series data and this constrain on the mean to variance ratio has been known to affect the goodness of the fit, since overdispersion is usually observed. We present a strategy for constructing a new class of overdispersed continuous time Markov counting processes with a nice interpretation in terms of adding noise to the rate parameters of standard compartmental models. As a motivating example, we consider a multistrain model for cholera incidence in Bangladesh.

Multilevel Modeling Technique and Bootstrap Variance Estimation in Longitudinal and Cross-Sectional Complex Survey Data Sets

✱Alomgir Hossain, University of Saskatchewan, Health Science Building, Community Health and Epidemiology, 107 Wiggins Road, Saskatoon, SK S7N5E5 Canada, *mdh512@mail.usask.ca*; Punam Pahwa, University of Saskatchewan; Bruce Reeder, University of Saskatchewan; Bonnie Janzen, University of Saskatchewan

Key Words: Multilevel technique, bootstrap, longitudinal, quasi-likelihood, pseudo-likelihood

Two most commonly used techniques to analyze longitudinal and cross-sectional complex survey data are: (i) standard regression techniques based on the quasi-likelihood for parameter estimation and the bootstrapping approach for variance estimation and (ii) multilevel technique based on pseudo maximum likelihood. Our study objectives are to explore the standard regression parameter estimation based on quasi-likelihood techniques, bootstrapping approach for variance estimation by using the longitudinal data from Canadian National Population Health Survey (NPHS) and weighted and un-weighted parameter estimation by using the cross-sectional Canadian Heart Health Datasets (CHHD), collected in1986–92, after controlling PSU level. The longitudinal NPHS began in 1994/95 with a sample size of 17,276 and data being collected every two years.

Correction of Bias from Nonrandom Missing Longitudinal Data

✱Cuiling Wang, Albert Einstein College of Medicine, 1300 Morris Park ave, Belfer 1303E, Bronx, NY 10461, *cuwang@aecom.yu.edu*; Charlie B. Hall, Albert Einstein College of Medicine

Key Words: missing data, MNAR, multivariate longitudinal data, multiple imputation, auxiliary data

A common problem in longitudinal studies is missing data. Likelihood-based mixed effects models give valid estimates when the data are missing at random, an untestable assumption without further information. In some studies, additional auxiliary information known to be correlated with the outcome might be available when the outcome of interest is missing; such additional data allows the MAR assumption to be tested and can be used to reduce or eliminate bias when the missing data process depends on the unobserved outcome only through the auxiliary information and the observed outcome. We apply three methods of utilizing the additional information: multivariate longitudinal analysis, multiple imputation, and using the auxiliary information as covariates. Simulations studies and analysis using telephone gait questionnaires to correct for missing longitudinal gait assessments were reported.

Modeling Change in Longitudinal Studies: Use Age Only or Initial Age and Time?

✱Christopher H. Morrell, Loyola College in Maryland, Mathematical Sciences Department, 4501 North Charles Street, Baltimore, MD 21210, *chm@loyola.edu*; Larry J. Brant, National Institute on Aging; Luigi Ferrucci, National Institute on Aging

Key Words: Multilevel modeling, Observational study, Recruitment bias, Regression

Since many epidemiological studies involve the study of individuals of different ages over time, it often becomes necessary to distinguish between and estimate both longitudinal and cross-sectional differences. This paper examines how the choice of age and time in modeling longitudinal data can affect the results. In particular, age can be decomposed into two components: age at entry into the study (first age) and follow-up time. The implication of using age or first age and time is described for a number of possible linear mixed-effects models that may be used to describe the longitudinal data. The two approaches are illustrated using a number of different examples of data taken from the Baltimore Longitudinal Study of Aging (BLSA). The examples illustrate that the added flexibility provided by the first age and time approach is usually necessary to adequately describe the data.

Comparison of Generalized Estimating Equations and Quadratic Inference Functions in Longitudinal Analysis of Data from the NLSCY Database

✱Adefowope O. Odueyungbo, McMaster University, Department of Clinical Epidemiology and Biostatistics, Hamilton, ON , *odueyuao@mcmaster.ca*; Dillon Browne, University of Guelph; Norri Akhtar-Danesh, McMaster University; Lehana Thabane, McMaster University

Key Words: Longitudinal analysis, Repeated responses, Marginal logistic models, Quadratic inference functions, Generalized estimating equations, Goodness-of-fit

GEE is often used to assess the effect of covariates on an outcome of interest in longitudinal/clustered data analyses. Parameter estimates obtained from GEE are consistent irrespective of the working correlation matrix. If the correlation matrix is mis-specified, estimates may become inefficient. Inferences from GEE may also be sensitive to outliers in observed data. In addition, GEE lacks an absolute measure of goodness-of-fit (GOF), which makes model selection a difficult task. QIF-an extension of the GEE-provides parameter estimates that are both consistent and efficient irrespective of the working correlation matrix. QIF also provides a GOF test with similar properties to the likelihood ratio test of generalized linear models. We compare results obtained from GEE and QIF, using dataset from NLSCY. The repeated response is a binary measure of hyperactivity-inattention in children.

59 Spatial and Space-Time Models and Applications

Section on Statistics and the Environment, WNAR

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Spatio-Temporal Modeling of Intra-Urban Variations in Air Pollution Concentrations

✱Adam A. Szpiro, University of Washington, Seattle, WA 98195-7232, *aszpiro@u.washington.edu*; Paul D. Sampson, University of Washington; Lianne Sheppard, University of Washington; Darren Wilton, University of Washington; Tim Larson, University of Washington; Thomas Lumley, University of Washington; Sara D. Adar, University of Washington

Key Words: Bayesian modeling, Spatio-temporal modeling, Air pollution

To estimate the effect of air pollution on cardiovascular health in a cohort study, it is necessary to predict individual exposure based on sparse measurements. The U.S. EPA has few monitors in any given city as the network is primarily designed for regional air pollution levels. To address variation near roads, the EPA-funded MESA Air project is carrying out additional monitoring. Two features of the data present challenges. First, to represent small-scale variation in traffic-related pollution, we must pay close attention to covariates and how they relate to local meteorology. Second, to take advantage of irregularly sampled data we need a rich statistical model with space-time interactions in the correlation. Estimation of the parameters depends on specialized computational techniques. We describe an approach to addressing the challenges outlined above and give illustrative results.

A Bayesian Spatial Model To Downscale Outputs from Numerical Models

✱Veronica J. Berrocal, U.S. Environmental Protection Agency, Research Triangle Park, NC , *berrocal.veronica@epamail.epa.gov*; Alan Gelfand, Duke University

In many environmental problems data arise from two sources: numerical models and monitoring networks. The first source provides predictions at the level of grid cells, while the second source gives measurements at points. Accommodating the spatial misalignment between the two types of data is of fundamental importance for both improved predictions of exposure as well as for evaluation and calibration of the numerical model. In this paper we propose a simple method to downscale the output from numerical models to point level. The model, specified within a Bayesian framework, regresses the observed data on the numerical model output using spatially varying coefficients, modeled as correlated spatial Gaussian processes. Moreover, the model can be easily extended to a dynamic implementation. As an example, we have applied our method to ozone concentration data for the Eastern US.

A Multivariate Spatial Model for Prediction of Storm Outages

✱Hongfei Li, IBM T. J. Watson Research Center, 1101 Kitchawan Road, Route 134, Room 35-238, Yorktown Heights, NY 10598, *liho@us.ibm.com*; Jonathan Hosking, IBM T. J. Watson Research Center

Key Words: Multivariate Spatial Model, Outage Prediction, Bayesian Hierarchical Model

We have developed a Bayesian hierarchical model for multivariate spatial data, with the aim of predicting the number of outages suffered by customers of an electric utility as a result of severe weather events such as high winds, thunderstorms, and tornadoes. Specific attention is given to building models that incorporate covariates as well as spatial correlation in a multivariate context. Bayesian inference is implemented by means of a MCMC algorithm.

An Investigation for the Trend of Global Warming Based on the Global Seawater Oxygen-18 Database from NASA

✱Qiqi Deng, The University of Minnesota, 1245 Fifield Ave, Saint Paul, MN 55108, *qiqi@stat.umn.edu*; Snigdhansu Chatterjee, The University of Minnesota

Key Words: Polynomial local regression, CUSUM

Background: NASA Goddard Institute for Space Studies (GISS) has the Global Seawater Oxygen-18 Database with over 22,000 observations recorded since about 1950, including date, ocean temperature, depth and coordinates, etc. In this study, we use this database to investigate the trend of global warming. Method: Polynomial local regression is applied to the data after appropriate transformations. Residual plots are checked for model validation. CUSUM chart and other statistical methods are used to detect the change of ocean temperature along years after adjusting the influences by other predictors, such as depth, spatial effects and seasonal effects. Results: The relationship between the ocean temperature and the predictors are complicated, and some interactions exist. The CUSUM chart suggests that there is an indication of slightly increasing trend for ocean temperature over years.

Probabilistic Assessment of Regional Climate Change by Ensemble Dressing

✱Christian Schoelzel, Meteorological Institute at the University of Bonn, Bonn, International 53121 Germany, *schoelzy@gmx.info*; Andreas Hense, Meteorological Institute at the University of Bonn

Key Words: climate change, ensembles, climate model, ensemble dressing, kernel dressing, Bayesian model averaging

We present an approach to obtain probabilistic information on climate change signals from ensembles of global and regional climate models. Since single-integration climate models only provide one possible realization of climate variability, ensembles are a promising way to estimate the uncertainty in climate modeling. Recently, several methods, mostly from the field of numerical weather prediction, have been developed (SKD, AKD, NGR, BMA). The approach presented here is related to kernel dressing. In opposition to BMA, the ensemble members are assumed to be indistinguishable. Hence, we attribute equal mixture weights while the dressing variance is extended to a spatio-temporal covariance matrix. The method has been applied to ensembles of the coupled general circulation model ECHO-G. Current applications include temperature and precipitation as well as seasonal variations over Europe.

Infill Asymptotic Properties of Tapered Maximum Likelihood Estimators

✱Juan Du, Michigan State University, Department of Statistics and Probability, MSU, A413 Wells Hall, East Lansing, MI 48824-1027, *dujuan@msu.edu*; Hao Zhang, Purdue University; V. S. Mandrekar, Michigan State University

Key Words: Covariance tapering, infill asymptotics, Matern covariance functions, maximum likelihood estimator

Massive spatial and spatio-temporal data are common these days, and the analysis of these massive data involves operations of a large covariance matrix. Repeatedly, inversion of the covariance matrix is needed in both the maximum likelihood and Bayesian inferences. When the sample size is extremely large, operations on the large covariance matrix are a numerical challenge. Covariance tapering is a technique to alleviate the numerical challenges. We investigate how the tapering affects the asymptotic efficiency of the maximum likelihood estimator (MLE) and establish the asymptotic properties, particularly the asymptotic distribution of the exact MLE and tapered MLE under the infill asymptotic framework for Matern model. We show that the tapered MLE is asymptotically as efficient as the true MLE for the Matern model under some tail conditions on the tapering function.

Optimal Monitoring of Plant and Animal Populations

✱David Marsh, National Center for Ecological Analysis and Synthesis, 735 State St., Suite 300, Santa Barbara, CA 93101, *marshd@wlu.edu*

Key Words: monitoring, population, conservation, ecology, environment, wildlife

Data from animal and plant monitoring programs form the basis for much environmental decisionmaking. Yet the design of monitoring programs is often based more on tradition than on rational analysis. I used a simulation approach to determine the optimal strategy for monitoring populations given a fixed sampling effort divided among many monitoring sites. Simulated populations were chosen to reflect high and low values for population parameters, as estimated from a time-series database. For most parameter sets, monitoring strategies that focused on monitoring more sites less frequently proved superior to approaches that concentrated monitoring effort on a few sites. Tests of the model on avian count data from the North American Breeding Bird Survey suggested that simulations could predict how well a given monitoring strategy would perform for all but the most variable populations.

60 Student Success in K–12 and Beyond ●

Section on Statistical Education

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Is NCLB Inflating Trends at Proficiency? A Cross-State Analysis of Trends Across Cut Scores

✱Katherine Furgol, The University of Iowa, Iowa City, IA 52245, *kfurgol@stat.uiowa.edu*

Key Words: interval censored data, nonparametric, No Child Left Behind, EM algorithm

As the pressure for states to increase the proportion of proficient students rises with the No Child Left Behind (NCLB) Act, incentives to teach low or high scoring students may diminish. We investigate this hypothesis using a cross-state dataset of proportions across multiple cut scores. The full paper develops and applies a nonparametric approach for establishing “expected” trends at different cut scores. A “normal-shift” model was developed to ascertain if more students than expected become proficient from Time 1 to Time 2. This model uses an EM-algorithm for estimating the parameters of normally distributed interval-censored data. Preliminary findings suggest that many states have 2003-2005 proficiency trends that are higher than expected; these findings will be augmented with 2007 state data. Conclusions include implications of these findings for NCLB’s academic equity provisions.

Data Analysis and Probability: An Examination of Criterion-Referenced Test Scores

✱Calli Holaway-Johnson, University of Arkansas, 340 N. West Ave. WAAX 302, Fayetteville, AR 72701, *cajohns@uark.edu*; Charles Stegman, University of Arkansas - NORMES; Sean Mulvenon, University of Arkansas

Key Words: education, data analysis, item analysis, test scores

Data analysis and probability content has been largely overlooked in K-12 school systems, as other content areas, such as algebra and geometry, have been considered to be more vital to student success. However, the National Council of Teachers of Mathematics (2006) has indicated that data analysis can be used to explore various other mathematical topics, such as proportions, fractions, estimation, place value, and number representations. This paper explored standardized criterion-referenced test scores for grades 3-8 on the Arkansas Benchmark exam Data Analysis & Probability (DAP) strand. Using data from 2005-2007, statewide trend analyses and item analyses were conducted, including an exploration of student performance disaggregated by Student Learning Expectation (SLE). Additional analyses that examined achievement gaps in DAP with regard to gender and ethnicity are also reported.

Improving Statistics Self-Efficacy: Learning from Errors and Feedback

✱T. Simin Hall, The University of North Carolina, Greensboro, 499 Reigate Drive, Kernersville, NC 27284, *tshall@uncg.edu*

Key Words: statistics word problems, self-efficacy, feedback, pattern recognition, calibration, on-line learning

Anxiety and low self-efficacy perceptions have been shown to affect students’ problemsolving ability. This research studied the impact of feedback on the changes in efficacy, and problemsolving scores for 138 students in two groups (feedback, no feedback) in an online learning environment. Students in the feedback group showed a statistically significant gain in their problem scores over the no feedback group; however, the mean efficacy scores were lower for both groups after the problemsolving experiment. The incongruence in problem scores with efficacy scores was attributed to students’ over rating of their abilities prior to actually performing the tasks. The process of calibration was identified as an explanation for the statistically significant correlation between problemsolving scores and post-efficacy scores for the feedback group.

The Relationship Between Learner-Centered Practices and Adolescent Achievement: A Latent Profile Analysis

✱Michelle Vanchu-Orosco, University of Denver, 4500 - 19th Street, #257, Boulder, CO 80304, *mvanchuo@du.edu*; Barbara L. McCombs, University of Denver Research Institute; Steven A. Culpepper, Metropolitan State College of Denver

Key Words: Multidimensional Scaling, Profile Analysis, Academic Achievement, Motivation, Teaching Practices

This study examined the effects of learner-centered practices on academic engagement and achievement via Profile Analysis using Multidimensional Scaling, extending original research by McCombs (2001, 2004). Using a national sample of fourth (n = 2188) and fifth (n = 2100) grade students, relationships between student motivational characteristics, perceived teaching practices, and goal orientation were examined. Multiple Regression analysis identified the portion of observed student achievement score accounted for by latent student profiles. Analyses indicated the presence of two profiles accounting for approximately 12% of variance in achievement. This supports previous research whereby students who believe their teacher provides a learner-centered environment are more likely to be self-efficacious, task mastery-oriented, and less likely to avoid work or require extrinsic rewards.

Who Gets Better Grades in the Statistics Course?

✱Jen-Ting Wang, SUNY-Oneonta, Oneonta, NY 13820, *wangj@oneonta.edu*; Shu-Yi Tu, University of Michigan, Flint

Key Words: Introductory Statistics, Correlation, Multiple Linear Regression

This study is to investigate the determinants of achievement in college introductory statistics course. Data concerning students' grades received in this course, the academic performance in high school and in college, as well as numbers of collegiate credits earned were collected from a public four-year liberal arts college. This study aims to identify the most significant factors of students' grades in this course. In addition, a comparison between performances of male and female students, as well as those of freshmen and non-freshmen was also examined. Class size effect was discussed as well. In addition to searching for the most important factors, the prediction model for the course grade was also established from multiple linear regressions.

Predicting Success in College Math Courses

✱Casey Gearheart, Roanoke College, 429 Ingal Blvd., Salem, VA 24153, *cngearheart@roanoke.edu*

Key Words: Statistics, Mathematics, Predictors, SAT Math Scores, Gender

The skills acquired in mathematics and statistics courses are becoming increasingly important in today's world of expanding technology. As a result, almost all college students are required to enroll in an introductory level mathematics or statistics course regardless of their major choice. This research project was designed to predict the characteristics of a student that best indicate whether the student will succeed or not in an introductory mathematics or statistics course. The study aims to improve placement of students and raise their success rate in these introductory classes. There were fifteen independent variables under study. Minitab© and SAS© statistical software packages were used to run general linear models and logistic regression models. The outcome variable, the final grade, was analyzed in two ways: as the letter grade and as pass/fail in the course.

A Causal Relationship Model of Factors Affecting Professional Competency Development of Computer Sciences Students

✱Wilailuk Seritrakul, Bangkok University, 9/1 Moo 5 Paholyothin Rd., Tambon Klong Nung, Amphur Klong Luang, Patum Thanee, 12120 Thailand, *wilailuk.s@bu.ac.th*

Key Words: Professional competency, Computer Science, Curriculum

The purpose of this research is to develop a causal relationship model of factors affecting professional competency development of Computer Science students. Six needs for this professional competency are concept of computer science, professional technique, context, intercommunication skill, integration skill, and adaptability. The population in this study were senior students from both public and private university located in Bangkok metropolitan areas. The results were: background of students, curriculum, in-class experience and out-of-class experience have direct effect on professional competency development but outside university reference group does not affect professional competency development directly. The background of students has an indirect effect on this development through outside university reference group, curriculum and both in class and out-of-class experience.

61 Statistical Software

Section on Statistical Computing, Section on Survey Research Methods

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Probabilistic and Fuzzy Matching as Applied to Record Linkage and Computerized Coding

✱Michael J. Wenzowski, Statistics Canada, R.H. Coats Building, Section 14-A, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *michael.wenzowski@statcan.ca*

Key Words: Record Linkage, Probabilistic Linkage, Fuzzy Matching, Automated Coding, Computerized Coding, Coding

As the overall speed of computer processing increases, the general perception of record linkage as an unwieldy and time-consuming process diminishes. Similarly, complex coding tasks which have traditionally been viewed as bordering on problematic are now seeing their implementation possibilities evolve towards the plausible. Each of these problem domains employ similar approaches to identifying the "best" candidate for either linkage or code assignment. This paper describes Statistics Canada's recent experiences in redeveloping their own generalized record linkage and computerized coding software, and highlights the convergence and associated overlap in techniques which has occurred across these two statistical processing domains. The overall system architecture employed, the techniques and algorithms implemented, and examples of the user interface developed are also presented.

Rank and Set Restrictions for Homogeneity Analysis in R

✱Patrick Mair, Wirtschaftuniversität Wien, *patrick.mair@wu-wien.ac.at*; Jan de Leeuw, University of California, Los Angeles

Key Words: Gifi-models, homogeneity analysis, nonlinear principal component analysis, nonlinear canonical correlation analysis, multiple correspondence analysis

The model family proposed by Gifi (1990) is a flexible framework for the analysis of multivariate data. The common properties shared by all Gifi-models are the specification of a loss function solved by ALS and transformations of the variables which lead to quantifications of the categories. The latter issue implies the concept of "optimal scaling" and allows to account for the variables scaling level. Starting from the basic model called homals (homogeneity analysis aka multiple CA) we present various extensions in terms of rank restrictions (nonlinear PCA) and restrictions on sets of variables (nonlinear canonical correlation analysis). We focus on recent methodological developments and on the R package "homals" (de Leeuw & Mair, 2008) which allows for the computation of these models and provides new visualization techniques of the results.

Sensitivity Analysis for Observational Studies: The ObsSens Package for R

✱Gregory L. Snow, Intermountain Healthcare, Eigth Avenue and C Street, Salt Lake City, UT 84143, *greg.snow@imail.org*

Key Words: Sensitivity analysis, Observational Studies, R packages

This paper introduces and demonstrates the obsSens package for R. It implements the sensitivity analysis methods for observational studies presented in the paper "Assessing the Sensitivity of Regression Results to Unmeasured Confounders in Observational Studies" by Lin, Psaty, and Kronmal. The package will be demonstrated using data from a study on platelet transfusions in premature babies. The study shows that a practice meant to help the newborns is unlikely to be beneficial and possibly harmful.

SAFAL: Statistical Analysis Functions Automating Language

✱Babubhai V. Shah, SAFAL Institute Inc., 22 Autumn Woods Dr., Durham, NC 27713, *babushah@earthlink.net*

Key Words: Statistical computing, Survey data analysis, Functional languge

The objective of this paper is to present an implementation of a prototype for SAFAL: Statistical Analysis Functions Automating Language. SAFAL uses a programming paradigm that is unique. Each statement of SAFAL is not an instruction to a computer, but is a definition of a symbol, using algebraic functional syntax. The statements in a SAFAL program are mathematical definitions (or input output specifications) of the relevant entities. Hence, the order in which they are specified in irrelevant. The design of SAFAL is aimed at statistical analysis in general, but the initial implementation as functions that more relevant to analysis of complex survey data. The basic data structure for SAFAL is a generalized multidimensional array (GMDA) with normal as well as virtual dimensions. The paper describes GMDA its underlying mathematics.

An Open Source Library for the Estimation and Evaluation of ACD Models

✱Guillaume Weisang, Bentley College, Morison 309, 175 Forest Street, Waltham, MA 02452, *gweisang@bentley.edu*

Key Words: Autoregressive Conditional Duration model, estimation procedure, tests statistics, open source

Following the publication of the first Autoregressive Conditional Duration (ACD) model by Engle and Russel (1998), numerous new specifications have been published in the literature. At the same time, tests statistics have been developed for the evaluation of ACD models. This paper provides an open source library of statistical procedures for the estimation of the standard linear ACD models, as well as the augmented models family presented by Fernandes and Grammig (2006). Source code of some statistical tests for the evaluation of the models, including distributional misspecification tests (Fernandes & Grammig 2005) and LM tests for misspecified mean equations (Meitz & Teräsvirta 2006), is provided as well. Several empirical applications are also presented.

Algorithmic Errors in the Estimation of Tobit II Models and the Corresponding Failure To Recognize Selection Bias

Thomas Zuehlke, Florida State University; ✱Anthony Kassekert, Florida State University, 655 Bellamy Building, Tallahassee, FL 32306-2250, *ajk05f@fsu.edu*

Key Words: Tobit II, Heckman Selection Model, Simultaneous Estimation, Monte Carlo Simulation, Software Review

Tobit II models (aka Heckman Selection models) are a standard statistical tool for detecting and correcting selection bias. ML estimation is complicated by the possibility of multiple roots to the score equations. Most software packages ignore this problem and may fail to converge to the global MLE even when consistent starting values are used. Convergence to the global MLE can be insured by use of a two-step algorithm which conducts a grid search over the bounded space of the error correlation, and then uses the conditional ML estimates as starting values for simultaneous estimation. The nature of the problem is illustrated using Monte Carlo simulation. Major software packages are then compared and found to suffer from the same algorithmic errors. Finally, replication of estimates for a sample of published data sets finds that roughly half of the studies report inaccurate estimates.

GLUMIP 2.0: SAS/IML Software for Planning Internal Pilots

✱John A. Kairalla, University of Florida, Division of Biostatistics, College of Medicine, PO Box 100177, Gainesville, FL 32610-0177, *jak@biostat.ufl.edu*; Christopher S. Coffey, The University of Alabama at Birmingham; Keith E. Muller, University of Florida

Key Words: sample size, re-estimation, power, adaptive

Internal pilot designs involve conducting interim power analysis (without interim data analysis) to modify the final sample size. Recently developed techniques have been described to avoid the type I error rate inflation inherent to unadjusted hypothesis tests, while still providing the advantages of an internal pilot design. We present GLUMIP 2.0, the latest version of our free SAS/IML software for planning internal pilot studies in the general linear univariate model (GLUM) framework. The new analytic forms incorporated into the updated software solve many problems inherent to current internal pilot techniques for linear models with Gaussian errors. Hence, the GLUMIP 2.0 software makes it easy to perform exact power analysis for internal pilots under the GLUM framework with independent Gaussian errors and fixed predictors.

62 Statistics in the Physical Sciences ●▲

Section on Physical and Engineering Sciences
Sunday, August 3, 4:00 p.m.–5:50 p.m.

Statistical Learning Methods for Neutron Transmission Tomography of Fuel Cells

✱Kevin J. Coakley, National Institute of Standards and Technology, Mailcode 898.03, 325 Broadway, Boulder, CO 80305-3328, *kevin@boulder.nist.gov*; Daniel S. Hussey, National Institute of Standards and Technology; Dominic F. Vecchia, National Institute of Standards and Technology

Key Words: cross-validation, fuel cell, hydrogen economy, neutron imaging, penalized likelihood, transmission tomography

In a proton exchange membrane fuel cell, effective water management is key to improved fuel cell performance and durability. Water enters as a humidified gas, and is produced by the oxidation reaction in the cathode.

Neutrons are readily scattered by water and by other materials in the fuel cell. For the dry and wet states of a fuel cell, there are spatially varying neutron attenuation images. Ideally, the difference of these attenuation images is proportional to the water density. We estimate a residual attenuation image from the wet and dry state projection data using penalized likelihood and penalized weighted least-square methods using a Huber penalty function with two adjustable regularization parameters that are determined by cross-validation. We reconstruct both simulated data and real data collected at the NIST Neutron Imaging Facility.

Determining Dispersion Effects in Cultured Fish

✱Tess Astatkie, Nova Scotia Agricultural College, Department of Engineering, P.O.Box 550, Truro, NS B2N 5E3 Canada, *tastatkie@nsac.ca*; Jim Duston, Nova Scotia Agricultural College

Key Words: Repeated measures, Arctic charr, Coefficient of variation, Range

Reducing variation (dispersion) in the size of intensively cultured fish is essential to enhance production efficiency and increase marketable product. Among salmon species, Arctic charr has the highest size variation, and this is a serous challenge for the aquaculture industry. Although large variation is an inherent trait for this species, there is a potential for reducing it by identifying factors with dispersion effects. However, almost all studies reported in mainstream aquaculture journals are on location effects only; and run a risk of promoting growth-enhancing treatments at the expense of increased variability. Dispersion effects are not reported mainly due to lack of experimental procedures applicable to untagged- and individually tagged-fish. In this paper, procedures for untagged and individually tagged Arctic charr, and how to determine dispersion effects will be presented.

Regression Analysis Investigation of the Link Between Hurricane Frequency and Climate Change

✱Roshanak Nateghi, Johns Hopkins University, Department of Geography and Environmental Engineer, Johns Hopkins University, Baltimore, MD 21218, *roshanak.nateghi@googlemail.com*; Steven Quiring, Texas A&M University; Seth D. Guikema, Johns Hopkins University

Key Words: count data regression, hurricane frequency, climate change, climatology

In this paper we analyze the possible link between hurricane counts in the Atlantic Coast and Gulf Coast regions of the United States and a number of measures of long-term climate variability and change. The region of interest is divided into four separate sections: Atlantic coast except Florida, Atlantic side of Florida, gulf side of Florida, and the rest of the Gulf. Regression models are implemented to analyze data that extends as far back as 1850. The analysis is done in an effort to study the frequency of hurricanes as a function of climate variables. The trend of hurricane occurrence over time is an indication of the impact of climate change on the frequency of hurricane landfalls in the above-mentioned regions. The results investigate whether global warming will indeed increase the number of hurricane landfalls on the east coast of the United States as some have predicted.

Modeling Spike Train Transformations in Brain Region

✱Dong Song, University of Southern California, *dsong@usc.edu*; Rosa H.M. Chan, University of Southern California; Vasilis Z. Marmarelis, University of Southern California; Robert E. Hampson, Wake Forest University; Sam A. Deadwyler, Wake Forest University; Theodore W. Berger, University of Southern California

Key Words: point processes, GLM, hippocampus, model selection, spike, Volterra kernel

In any brain region (e.g., hippocampus) information is represented in the ensemble firing of populations of neurons (i.e., spikes trains). Brain regions process the information by transforming the incoming spike trains into outgoing spike trains. We present a physiologically plausible model to describe such nonlinear dynamic transformation. This model consists of a set of feedforward Volterra kernels, a feedback Volterra kernel, a threshold and a noise term. Assuming Gaussian distribution of the noise, this model can be formulated as a generalized linear model with a probit link function. Model selection techniques such as information criteria and cross-validation can be used to reduce the complexity of the Volterra kernels. We demonstrate that the model can sufficiently capture the hippocampal CA3-CA1 spike train transformation in rats performing a memory-dependent behavioral task.

Spam Email Detection and Classification: A Bayes/Empirical Bayes Approach

✱Cheng C. Chen, Texas A&M University, MSC 172, University Blvd, Mathematics Department, Kingsville, TX 78363, *kfccc01@tamuk.edu*; Thomas Chen, Southern Methodist University; Rajat Sethi, Texas A&M Health Science Center

Key Words: SPAM, Bayes, Empirical Bayes, Classification, Calssification Error Rate, Naive Bayes

Spam filtering technologies based on statistical methods are not new such as Naive Bayes classifier However, there is a flawed assumption underneath the Naive Bayes approach: The probability of two or more spam evidences will result in less probability than any single spam key word. The reason underlying this flaw is the intent of reducing the content of past spam and updating activities and hence CPU processing power. In this paper, we propose a content based spam detection and classification method based on Bayes/Empirical Bayes approach. We claim that the required CPU power and memory are no more than the Naive Bayes classifier without the above mentioned technical flaw of Naive Bayes approach. The error classification probability is also discussed. The application of desease classifications based on pathogensis measurements will also be discussed.

Estimating the Skewness and Kurtosis of Nonlinear Time Series: Analysis of Turbulent Flows with Coherent Structures

✱Alexander Gluhovsky, Purdue University, Department of Statistics, 250 N. University Street, West Lafayette, IN 47907-2066, *aglu@purdue.edu*

Key Words: Nonlinear time series, Subsampling, Turbulence

Turbulent flows are characterized by non-Gaussian statistics exposed by the values of skewness and kurtosis. Nonzero skewness and increased values of kurtosis in observed time series are attributed to the occurrence of coherent structures. Due to the major role these structures play in turbulence, skewness and kurtosis are increasingly investigated, however, with a questionable accuracy of their estimates. In the talk, a subsampling-based methodology will be described for obtaining reliable confidence intervals for the skewness and kurtosis from observed and modeled data, with applications to the analysis of atmospheric time series. This work is supported by the National Science Foundation Grants ATM-0514674 and ATM-0541491.

A Comparative Study of Mechanistic and Statistical Modeling of Water Quality Parameters in Hillsborough Bay, Florida

✱Keith E. Hackett, University of South Florida, 4876 28th Street South, St Petersburg, FL 33712-4341, *khackett2@tampabay.rr.com*; Chris P. Tsokos, University of South Florida

Key Words: Hillsborough Bay, water quality, EFDC, statistical modeling, model calibration, mechanistic modeling

Significant improvements in the water quality of Hillsborough Bay, Florida have been documented in the recent past as a result of the concerted efforts of the Nitrogen Management Consortium under the auspices of the Tampa Bay Estuary Program. With the projected increase in human population living in the watershed, it is important to identify the most cost-effective methods to control loadings and continue to improve water quality in the bay. Two modeling techniques, mechanistic and Bayesian, have been employed in order to better predict the future water quality in Hillsborough Bay based on projected increases in population. A comparison of the two modeling techniques reveals the strengths and weaknesses of the methods. In addition, this paper discusses methods of using Bayesian statistical modeling techniques to improve the calibration of mechanistic models in general.

63 Semiparametric Inference Under Partially Linear, Single-Index, and Related Models ●▲

Section on Nonparametric Statistics, IMS

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Partial Linear Regression Model Checking with Berkson Measurement Errors

✱Weixing Song, Kansas State University, 2220 Westchester Park, Apt.10, Manhattan, KS 66503, *weixing@ksu.edu*

Key Words: Marked Empirical Process, Asymptotic Distribution Free, Brownian Motion

This paper discusses asymptotically distribution free tests for the lack-of-fit of a partially linear regression model when both covariates in the parametric and nonparametric parts are subject to the Berkson measurement error. These tests are based on a martingale transform of a certain marked empirical process of calibrated residuals. Under the null hypothesis, these test statistics weakly converges to a distribution related to the maxima of the Brownian motion. Consistency of the tests under some fixed alternatives and the asymptotic power under some local alternatives are also discussed. Simulation studies are conducted to assess the performance of the proposed test.

On Estimation of Censored Single-Index Models

✱Zhulin He, University of Calgary, Department of Mathematics and Statistics, Calgary, AB T2N 1N4 Canada, *zhulinhe@math.ucalgary.ca*; Xuewen Lu, University of Calgary

Key Words: AFT model, Censored single-index model, MWAVE, OPWG, Kaplan-Meier weights

The accelerated failure time model assumes a linear relationship between event time and covariates. We propose a censored single-index model to extend the identity link function in the AFT model to an unknown link function. We provide two estimation methods for it – minimum weighted average conditional variance estimation (MWAVE) and outer product of weighted gradients (OPWG). These methods use the Kaplan-Meier weights to account for censoring. The MWAVE method estimates parameters by minimizing the overall weighted approximation errors which are calculated from the response variable and the estimated link function. The OPWG method works on the eigenvector corresponding to the largest eigenvalue from the outer product of weighted gradients of the estimated link function. Simulation studies and real data examples are used to illustrate the proposed methods.

Bandwidth Selection and Goodness-of-Fit Tests for Single-Index Models

✱Wei Lin, The Ohio State University, Department of Mathematics, 321 Morton Hall, Ohio University, Athens, OH 45701, *linw@ohio.edu*

Key Words: Single index model, Partially linear single index model, bandwidth selection, goodness-of-fit test

Single-index models (SIMs) have received much attention in the literature due to the ability in dimension reduction. However, the bandwidth selection problem for the kernel estimation of the regression function has been addressed only under very restricted conditions that can hardly be satisfied in practice. In this talk, I will provide a proof for the asymptotic normality of the LSE of the index vector in a SIM, using a practically meaningful data-driven bandwidth selection method, assuming conditions much weaker than those in the literature. The result is then extended to the class of Partially Linear SIM. Moreover, I will introduce two simple goodness-of-fit tests for linear models against SIMs. The asymptotic normality of the TS will be established and the simulation results will be presented to show the finite sample performance of the proposed tests.

Penalized Quadratic Inference Functions for Single-Index Models with Longitudinal Data

✱Yang Bai, The University of Hong Kong, Room 518, Meng Wah complex, Hong Kong, University, Hong Kong, Hong Kong, CA China, *baiyangg@hkusua.hku.hk*

Key Words: longitudinal data, penalized splines, quadratic inference functions, single-index model

In this paper, we focus on single-index models for longitudinal data. We propose a procedure to estimate the single-index component and the unknown link function based on the combination of the penalized splines and quadratic inference functions. It is shown that the proposed estimation method has good asymptotic properties. We also evaluate the finite sample performance of the proposed method via Monte Carlo simulation studies. Furthermore, the proposed method is illustrated in the analysis of a real data set.

Testing for Linearity of the Nonparametric Component of a Partially Linear Model

✱Chin-Shang Li, University of California, Davis, Division of Biostatistics, Room 145, Department of Public Health Sciences, Davis, CA 95616, *cssli@ucdavis.edu*

Key Words: B-splines, Partially linear model, penalized least-squares, Wald-type test

A linear combination of fixed knot cubic B-splines is used to estimate the nonparametric component of a partially linear model. We estimate the model parameters by minimizing the penalized least-squares criterion. A Wald-type test based on the proposed estimation method is proposed to assess the linearity of the nonparametric component. We study the power performance of the proposed test through simulations and illustrate its practical use with a real-life example.

64 Modeling of Brain Imaging and Related Data

Biometrics Section, Section on Nonparametric Statistics, Biopharmaceutical Section, WNAR

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Wavelet Estimation and fMRI Applications

✱Zibonele A. Valdez-Jasso, University of Texas at Dallas, 7760 McCallum Blvd Apt 20317, Dallas, TX 75252, *zav061000@utdallas.edu*; Sam Efromovich, University of Texas at Dallas

Key Words: Wavelets, fMRI, Denoising, Wavelet Estimation, Correlated Noise

The advance in medical research has lead to a great number of new techniques and data registration that creates an ocean of opportunities for statisticians to test and create new methods. Ultra frequent fMRI techniques is one of the best scenarios for the application and improvement of wavelets ideas. Crucial information can be obtained from those readings through the application of wavelets, therefore the best estimation and denoising methods are essential to enable reliable statistical results. An improved Wavelet Denoising method in the presence of correlated observations will be discussed and results of its application to both simulated and real data shown.

Thresholding Complex Magnetic Resonance Images Using Magnitude and Phase

✱Daniel B. Rowe, Medical College of Wisconsin, Department of Biophysics, 8701 Watertown Plank Road, Milwaukee, WI 53226, *dbrowe@mcw.edu*

Key Words: MRI, complex, thresholding, phase, correlated chi-square, voxel

It is well known that MRIs contain signal in the real and imaginary channels. It is desirable to separate voxels that contain tissue signal and noise from those that contain purely noise. Generally this separation (thresholding) utilizes only the magnitude portion of the images. A rigorous statistical method is presented that utilizes both the magnitude and phase to threshold image voxels. Using the bivariate normality of the real and imaginary voxel measurements with phase coupled means, a likelihood ratio statistic is derived. This likelihood ratio statistic simplifies to a familiar form that is F-distributed in large samples. In small samples, Monte Carlo critical values must be used for thresholding due to correlated chi-square statistics in the numerator and denominator of F. This method is successfully applied and shown to result in increased image contrast.

Estimating Distributions of Onset Times and Durations from Multisubject fMRI Studies

✱Lucy F. Robinson, Columbia University, 1255 Amsterdam ave. room 1005, MC 4409, New York, NY 10027, *lucy@stat.columbia.edu*; Martin Lindner, Columbia University

Key Words: fMRI, brain, change point, spatial

Most statistical analyses of fMRI data assume that the timing of brain activity is known a priori. As an alternative, we propose a technique for estimating voxel-specific distributions of activation onset and duration from the fMRI response, explicitly modeling each subject's onset and duration as random variables drawn from unknown distributions. These distributions are estimated assuming no functional form, along with the probability of activation at each time point. A hidden Markov random field model incorporating structural information about the brain is used to segment the brain image based on information from onset and duration profiles. We apply this method to an fMRI study of a single induction of state anxiety.

A Social Network Probability Map of the Brain

✱Lynne Seymour, The University of Georgia, Department of Statistics, The University of Georgia, Athens, GA 30602, *seymour@stat.uga.edu*

Key Words: Brain Imaging, Social Nework Model, Principal Components, 2-D White Noise, Ising Model

We build a probability map of the brain based on fMRI data from multiple subjects using a social network model. The neighborhood system for the social network is built via principal components breakdown of the image. A test for two-dimensional white noise based on the Ising model is presented as a means for choosing which principal components to remove from the construction.

Identifying Active Regions of the Brain via Social Networks Modeling in Single Subjects

Lynne Seymour, The University of Georgia; ✱Ana M. Bargo, The University of Georgia, Department of Statistics, The University of Georgia, Athens, GA 30602, *amoura@uga.edu*; Abhyuday Mandal, The University of Georgia

Key Words: Pseudo-likelihood, Social Networks, MCMC, Gibbs sampling, fMRI

Functional Magnetic Resonance Imaging is a relatively new and noninvasive procedure that uses magnetic resonance imaging techniques to measure metabolic changes that result in activation in a region of the brain. In this paper, we propose an innovative technique, which utilizes a social network model, to predict the active brain voxels for single subjects. Bayesian data analysis techniques are employed to identify the structure of the social network. The parameters of the associated models are estimated by Pseudolikelihood and MCMC methods, both of which lead to consistent estimates of activation probabilities.

Adapting FDR Thresholding Procedures to Neuroimaging

✱Lynn E. Eberly, The University of Minnesota, 420 Delaware St SE, MMC 303, Minneapolis, MN 55455, *lynn@biostat.umn.edu*; Brian S. Caffo, Johns Hopkins University

Key Words: medical imaging, false discovery rate, spatial correlation, thresholding

In functional neuroimaging, voxels in a brain map may represent clinical contrasts: activation in disease versus control. Conclusions are based on thresholding: large voxel values are declared significant. Such thresholding is also done in gene expression (detect elevated expression), discrete wavelet transforms (select wavelet coefficients), and other areas. Such procedures are based on choosing a threshold so that the False Discovery Rate (FDR) is controlled. Because of the inherent spatial nature of neuroimaging, procedures developed in other contexts may not perform well here. We compared the Storey procedure with other (semi)parametric procedures developed for other areas, and with AIC and BIC procedures. We compared FDR, estimated FDR, false nondiscovery rate (FNR), power, and MSE. We found surprisingly good results for BIC, poor for parametric, and widely varying MSE.

65 Credit; Bootstrap

Business and Economics Statistics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Predicting Credit Card Attrition Curves with Survival Models

✱Alfred Furth, South Dakota State University, 109 N Detroit Ave, Sioux Falls, SD 57110, *alfred.furth@sdstate.edu*

Key Words: Cox Regression, Accelerated Failure Time, Risk Models, time to event, survival

Portfolio attrition rates are important analytics monitored by a credit card company's finance and risk departments. Accurately predicting future attrition rates is important to a company's profitability. Cardholders contribute to the attrition rate in two ways, they can close their account or they can stop making payments and charge-off. Prediction of closure and charge-off rates can be framed as simple time to event problems. I will discuss the application of Cox proportional hazard (COX) and accelerated failure time (AFT) regression models to independently predict closure and charge-off rates with a set of risk scores. However, the independence assumption used in computing charge-off and closure rates via COX and AFT models is not completely true. Cardholder accounts have a risk of charging-off and closing. A competing risks model to account for this dependency will also be fit.

Modeling the Dynamic of Consumer Credit Utilization

✱Jie Chen, Bank of America, N.A., 201 N Tryon Street, Charlotte, NC 28202, *jie.chen@bankofamerica.com*; Agus Sudjianto, Bank of America, N.A.; Runze Li, The Pennsylvania State University; Richard Liu, Bank of America, N.A.

Key Words: growth, consumer credit, differential equation, latent process, dynamics

We model the dynamics of consumer credit utilization as a mixture of two latent processes representing unobserved underlying consumer behaviors. The latent processes are represented by a set of differential equations to capture two distinct growth characteristics of credit utilization. The model is also expanded to capture the effect of time-varying exogenous economic covariates. The framework is applied to a large portfolio of consumer credit. The model goodness of fit and its parameter estimation properties are investigated.

Frequency Domain Bootstrap Test for Time Series Linearity

✱Jane L. Harvill, Baylor University, Department of Statistical Science, One Bear Place, #97140, Waco, TX 76798-7140, *jane_harvill@baylor.edu*

Key Words: Nonlinear time series, Bispectral density function, Bootstrap

Existing frequency domain tests are based on approximations and asymptotic distributions of the sampling distribution of the estimated normalized bispectrum. Even when the time series is reasonably long, it has been shown that the distributional properties may not be precise in terms of type I error rates. The tests are also known to suffer from lack of power. We present a bootstrap version of a test for the Gaussianity and linearity of a sixth-order stationary time series. The size and power of the test is compared to popular time domain and frequency domain tests.

Generating Random Integer Sequences for Marketing and Risk Analytics

✱Michial Thompson, Strounine Thompson Technology Group, PO BOX 990934, Boston, MA 02199-0934, *mt@sttechgroup.com*

Key Words: random, randomization, marketing, risk management

Consider (1) We do a test by randomly assigning products to approved applicants. An applicant that submits multiple applications or cosigns on multiple applications must get the same random product assignment each time. (2) We must select a random sample from a dataset of 60M credit card transactions such that if one transaction tied to an account is selected randomly then every transaction tied to that account must be selected randomly. Our methodology solves problems like those above while meeting the requirements of repeatability, platform-independence and speed. We do so while reading the dataset only once and without the need for an external list of previously selected records. For example, this can be done in a simple data step in SAS by opening only one dataset and creating only one data set. These requirements can not all be met by built-in randomization routines.

Pricing Interest Rate Options in a LIBOR Market Model

✱Ting-Pin Wu, National Taipei University, 67, Sec. 3, Ming-shen E. Rd., Taipei, 104 Taiwan, Taipei, 104 Taiwan, *tpwu@mail.ntpu.edu.tw*; Shih-Kuei Lin, National University of Kaohsiung

Key Words: LIBOR Market Model, Martingale Pricing Method, Interest Rate Exchange Options

This paper presents a new approach to approximate the distribution of a forward swap rate under the LIBOR market model and then employs it to price derivatives involved with constant-maturity swap rates. We use this model to price LIBOR vs. swap interest rate exchange options, which have not been formally examined in the previous literature under a multifactor LIBOR market model to our best knowledge. The numerical examples show that the approximate pricing formulas are sufficiently accurate as compared with Monte Carlo simulation.

On the Term Structure of Credit Risk Spreads Under Levy Processes

✱Budhi A. Surya, Bank of America, N.A., Blk 9 Raffles Place #18-00 Republic Plaza, Global Quantitative Risk Management, Singapore 048619, S 048619 Singapore, *budhi.a.surya@bankofamerica.com*

Key Words: Credit risk, Credit spreads, Endogenous default, Optimal capital structures, Levy processes, First-passage-time problems

The determination of credit spreads has in fact been the ultimate goal of most credit risk models. From the structural models of Black and Scholes (1973), Merton (1974), and Leland and Toft (1996), one can deduce that the term structure of credit spreads should always have zero values at zero. Empirical study such as found in Sarig and Warga (1989) suggests disappointing results that the actual credit spreads do not have such feature. This drawback is due to the fact that the models have relied heavily on diffusion processes to model the evolution of firm value. Quite recently, a class of jump-diffusion processes have been used as an alternative to their models, see Zhou (1996), Hilberink and Rogers (2002), and Chen and Kou (2007). In this paper we extend these models for a general Levy process, and provide an estimate and sufficient conditions for the positivity of the credit spreads.

66 Statistical Methods for Genome-Wide Association Studies

Biometrics Section, Section on Statistics in Epidemiology

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Genotyping and Inflated Type I Error Rate in Genome-Wide Association Case/Control Studies

✱Joshua N. Sampson, Yale University, Center for Statistical Genetics and Proteomics, 300 George St, Suite 503, New Haven, CT 06511, *joshua.sampson@yale.edu*; Hongyu Zhao, Yale University School of Medicine

Key Words: genotyping, association, GWA, calling

The empirical distribution of the test statistic from a case/control genome-wide association study rarely resembles the predicted null distribution. The result is a high type I error rate. One cause is that the distribution of the intensity measurements (i.e., sum and contrast) may not completely separate for two of the genotypes. We examine the peculiar effects of genotyping and how genotyping can inflate the type I error rate for SNPs. Currently, methods assign genotypes using either the Mahalanobis distance (MD) or MLE. We show that the called genotypes for two subjects can be correlated even when their true genotypes are independent, and that this dependency can inflate type I error. We also show that the estimated genotype proportions need not converge to the true population proportions, and that this inconsistency can be corrected by assigning distances based on a weighted MD.

Power of Model Selection Methods for High-Dimensional, Genome-Wide Association Data

✱Zheyang Wu, Yale University School of Medicine, Department of Epidemiology and Public Health, Suite 503, 300 George St., New Haven, CT 06511, *zheyang.wu@yale.edu*; Hongyu Zhao, Yale University School of Medicine

Key Words: power, model selection, high dimensional data, genome-wide association study, SNP

Genome-wide association studies (GWAS) are characterized by a large number of SNPs and potential interactions among SNPs to affect disease susceptibility. Single SNP-based analysis is not optimal in the presence of interactions. Exhaustive search among models involving multiple SNPs is computationally intensive, and it also suffers from an increased chance of finding false positives due to fast increased number of models explored. Forward selection method may miss the SNPs with small marginal effects yet large interaction effects. A delicate balance needs to be achieved between computational efficiency, statistical power, and low false positive rates. This problem has been partly explored by limited simulation studies. Here we show to analytically calculate the statistical power of different SNP selection methods, and address various factors that have impact on more appropriate choices.

Bayesian Inference of Epistatic Interactions with Applications to HIV Drug Resistance

✱Jing Maria Zhang, Harvard University, 1 Oxford Street, Statistics Department, Cambridge, MA 02138, *zhang3@fas.harvard.edu*; Jun S. Liu, Harvard University

Key Words: Bayesian inference, epistatic interaction, HIV drug resistance, Bayesian network

Recently, a Bayesian Epistasis Association Mapping (BEAM) method for large-scale case-control genetic studies was proposed and shown to be significantly more powerful to detect an interaction than existing approaches. Here we applied this method to HIV-1 drug resistance data, found multiple drug resistance positions. Our findings are highly consistence with IAS-USA drug resistance mutation list (2007 update). Surprisingly, most of the mutations on the list are contributing jointly rather than individually. An independent database, Genotype-phenotype database, offers confirmation for our findings. In order to look at detailed structure of drug resistance, we use Bayesian network to fit the data. We proposed a new recursive Bayesian method to learn the structure of BN. It can be shown that the method asymptotically identifies the perfect map, or an inclusion-optimal map.

Dimensionality Reduction for Genetic Background Estimation in Whole-Genome Association Studies

✱Jasmin Divers, Wake Forest University Health Sciences, WC-2326, Medical Center Blvd, Winston-Salem, NC 27107, *jdivers@wfubmc.edu*; Jennifer B. Erway, Wake Forest University; Matthew L. Stiegert, Wake Forest University Health Sciences; Carl D. Langefeld, Wake Forest University Health Sciences

Key Words: GWAS, Population stratification, Type I error, Power, Random Projections, Principal Components, Measurement error

Genome-wide association studies (GWAS) have successfully identified genetic variations that predict risk of disease. Population stratification and admixture may confound genetic association tests. Principal component (PC) analysis has been utilized to control for these confounding influences on both power and type I error in population-based studies. Here, PCs are selected from a dispersion matrix computed from the genotyped data and used as covariates in linear or semiparametric models testing for association. However, directly computing the PCs becomes impractical as the number of markers and sample size increase. We study the performance of different formulations of the Lanczos algorithm and random projection methods. We evaluate their precision and their ability to provide adequate control in GWAS involving individuals with different ancestral backgrounds.

Determining Heritability Using Second-Level Testing for Genome-Wide Association Studies

✱Jae Brodsky, University of California, Los Angeles, 8125 Math Sciences Bldg., Department of Statistics, Los Angeles, CA 90095-1554, *brodsky@ucla.edu*

Key Words: genetics, genome-wide association studies, multiple comparisons, data reduction

Although genome-wide association studies (GWAS) have been used to successfully map loci associated with physical traits of interest, in many cases the results contain fewer SNPs than expected or even no SNPs at all. This does not necessarily indicate that there is no genetic association: SNPs with weaker associations or SNPs which are not independently associated with the phenotype will be missed at the genome-wide level. Second-level testing methods can be used to determine if there are signals in the data which cause the distribution of test statistics or p-values to be different from their expected null distribution. We apply such a method to a large GWAS and found results which are consistent with previously performed studies. We also discuss modifications to handle genetic data, which does not satisfy the conditions for optimality.

A Score Statistic for Testing for Genetic Association Given Linkage

✱Jeanine Houwing-Duistermaat, LUMC, Dept of Medical Statistics, S-5-P, PO Box9600, Leiden, International 2300 RC Leiden, *j.j.houwing@lumc.nl*; Hae-Won Uh, LUMC; Rick van Minkelen, LUMC; Marieke de Visser, LUMC

Key Words: genetic association, genetic linkage, score statistic, thrombosis, statistical genetics

Genome-wide linkage studies are often followed by association studies of candidate genes located under linkage peaks. Single-nucleotide polymorphisms (SNPs) are typed in a set of controls and in the affected relative pairs of the linkage study. To model this type of data, we extend the model for linkage of Kong and Cox (1997) with an association term. This model appears to correspond to a relative risk model with the unknown genetic factor as exposure. We derive a score statistic to test for association given linkage for both autosomal SNPs as well as x-linked SNPs. To illustrate the statistic we analyze data from an affected sibling pair linkage study on Thrombosis (209 families). Genotypes of SNPs located under a peak at one of the autosomal chromosomes and under a peak at the X chromosome are available in the affected sibling pairs and in 331 controls.

What's the Best Statistic for a Simple Test of Genetic Association in a Case-Control Study?

✱Chia-Ling Kuo, University of Pittsburgh, 5700 Centre Ave Apt 805, Pittsburgh, PA 15206, *chialing.kuo@gmail.com*; Eleanor Feingold, University of Pittsburgh

Key Words: genetic association test, case-control study, chi-squared test, trend test, logistic regression, genome-wide association study

Genomewide genetic association studies typically start with univariate statistical tests. In a case-control study, one can use a 1 df allele-based test, a 1 df or 2 df genotype-based test, or a compound procedure that combines two or more of these statistics. While there are a number of statistical papers that make power comparisons among subsets of these methods, none has comprehensively tackled the question of which method is best for a genome scan. In this paper we first compare the power of the different procedures to detect a single locus and then address whether or when it is a good idea to include covariates. Finally, we consider the performance of the statistics in the context of a genome scan.

67 Adaptive Design in Clinical Trials ●▲

Biopharmaceutical Section, WNAR, Biometrics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Adaptive Randomization: The Preferred Randomization in Multicenter Clinical Trials

✱Bob Zhong, Centocor, Inc, 200 Great Valley Parkway, Malvern, PA 19355, *bzhong@cntus.jnj.com*; Lilianne Kim, Centocor, Inc. ; Youyi Shu, Centocor, Inc.; Yong-Cheng Wang, Biogen Idec, Inc.

Key Words: clinical trials, adaptive randomization, balance

Therearemanyrandomizationmethodsthathavebeenusedtoallocatesubjects to treatment groups in clinical trials. The commonly seen randomization methods are complete (simple) randomization, stratified permuted block randomization and dynamic randomization (the minimization procedure with biased coin assignment proposed by Pocock and Simon (1975)). The advantages and disadvantages of the three commonly used randomization methods are evaluated for balance (treatment and prognostic factor levels) and statistical analysis methods. Since the adaptive randomization has more advantages over other two-randomization methods, we conclude that adaptive randomization is the preferred method of randomization. The preferred statistical analysis methods are also discussed and recommended.

A Hybrid Adaptive Design for Phase 2 Clinical Trials

✱Richard J. McNally, Pharmion Corporation, 9900 W. 109th St., Suite 300, Overland Park, KS 66210-1438, *rmcnally@pharmion.com*

Key Words: Clinical Trials, Adaptive Designs, Bayesian, Stopping Rules, Futility Analysis

A design for a randomized adaptive 3 arm study is presented where the primary objective is to determine what arms are suitable for further study. The main adaptive features are 2 sets of stopping rules: the first drops an arm if the absolute response rate is below a given target, and the second drops an arm in which the response rate lags behind the response rate in the best performing arm. These adaptive rules are derived in a manner that resembles the criteria used in Bayesian adaptive design (cf. Wathen and Cook), but the criteria are set in advance as the number of responses. Thus, the adaptive design can be regarded as a hybrid of Bayesian and frequentist. Derivation of the stopping rules and the operating characteristics of the hybrid design are discussed. Discussion of the latter includes in situations where there are slight deviations from the planned design.

An Integrated Algorithm for Adaptive Design: Sample Size Re-Estimation, Interim Treatment Selection, and Futility—Method and Computing Package

✱Yi He, sanofi-aventis, 1402 Fox Run Dr., Plainsboro, NJ 08536, *hydinghua@gmail.com*; Zhenming Shun, sanofi-aventis; Martin Roessner, sanofi-aventis

Key Words: Adaptive Design, Sample Size Re-estimation, Stopping Boundary, Treatment selection, Type I error, Interim Analysis

Sample size re-estimation based on observed treatment difference at interim can ensure adequate power and potentially save a large amount of time and resources in clinical trials. The type I error and stopping boundaries were mathematically quantified according to a broad set of scenarios which are practical considerations in conducting clinical trials. If proper adjusting rules are used, the inflation of type I error can be well controlled. The proposed stopping boundary combines the advantages of the group sequential and sample size re-estimation methods and is more efficient than either one alone. The proposed method was implemented in stand-alone software AdaptDes, with user-friendly interfaces. A demonstration of the software will be given during the talk.

Hypothesis Testing in Adaptive Design for Drop-the-Losers

✱Gang Li, Johnson & Johnson Pharmaceutical R&D, LLC, 920 Route 202, P.O. Box 300, Raritan, NJ 08869, *gli@prdus.jnj.com*; Yining Wang, Johnson & Johnson Pharmaceutical R&D, LLC; Peter Ouyang, Johnson & Johnson Pharmaceutical R&D, LLC; Gorden Lan, Johnson & Johnson Pharmaceutical R&D, LLC.

Key Words: Adaptive design, power function

Adaptive design for drop-the-losers combines two sequential studies into one study with two stages. We consider a study which starts with several experimental treatment groups and a control group. At the end of Stage 1 (interim), the enrollment of patients into inferior treatment groups (losers) will be discontinued. In Stage 2, we only continue to recruit and randomize patients into the control group and two superior treatment groups (winners) in Stage 1. At the end, data from Stages 1 and 2 will be pooled to compare the two treatment groups with the control group. Tests of likelihood ratio type are obtained when data are normally distributed. These procedures and several other approaches will be evaluated and compared through simulations.

Comparison of the Weighted and Unweighted Z-Statistics Used in a Two-Stage Adaptive Clinical Trial After Sample Size Change

✱John Ouyang, Otsuka Pharmaceutical Development & Commercialization, Inc., 2440 Research Blvd, Rockville, MD 20850, *john.ouyang@otsuka.com*; Peter H. Zhang, Otsuka Pharmaceutical Development & Commercialization, Inc.

Key Words: Type I error, Sample size re-calculation, Adaptive design, Conditional power, weighted z-statistic

The weighted z-statistic was proposed by Cui, Hung, and Wang (1999) to maintain the type I error in sample size increase after an interim analysis. It is proved that the weighted z-statistic is uniformly more powerful than the unweighted z-statistic, if the sample size is increased in a two-stage adaptive study when the condition of maintaining the type I error of the unweighted z-statistic given by Chen, DeMets, and Lan (2004) is met. It is also proved no inflation in type I error in the weighted z-statistic when sample size is changed. Use of the weighted z-statistic in two-stage adaptive design provides flexibility and efficiency. When it is used, it is recommended that the timing of the analysis of the 1st stage data would be set relatively early, if the parameter assumption has strong support; or set relative late, if the parameter assumption is made without much data support.

Seamless Phase II/III Clinical Trials Using the Likelihood Approach

✱Guijun Yang, Tianjin University of Finance and Economics, Department of Statistics, No. 25, Zhujiang Road, Hexi District, Tianjin, 300222 P.R. China, *tjyangguij@yahoo.com.cn*; Samuel S. Wu, University of Florida; Mark C.K. Yang, University of Florida; Anqi Sun, University of Florida

Key Words: Seamless clinical trial, likelihood ratio test, type I error, two-stage design, treatment selection, power

Seamless designs that combine the phase II dose finding trial and the confirmatory phase III trial have attracted considerable interest for its ability to improve statistical efficiency through joint analysis of data and to eliminate the gap between the two phases. Bretz et al. (2006) and Wang (2006) have proposed two-stage designs using the p-value combination and the conditional error function approaches. In this paper, we consider the combination of the two phases using the sequential likelihood ratio test. Our method extends the procedures of Thall, Simon and Ellenberg (1988) by allowing rejection of the null hypothesis at the time of selection or later (possibly more interim) analysis. We show that the type I error is controlled by those under the global null hypothesis. It can also be shown that likelihood approach has better power for a large class of alternatives.

Adaptive Designs for Dose-Ranging Phase I/II Trials

✱Naum M. Khutoryansky, Novo Nordisk Inc, 100 College Rd. West, Princeton, NJ 08540, *nakh@novonordisk.com*

Key Words: Adaptive design, Fisher's combination test, Dose ranging, Dose escalation, Minimum effective dose

Adaptive designs are described for phase I/II clinical trials with objectives to target the minimum effective dose, and/or to find the dose response curve and therapeutic range. The following adaptations are under consideration: sample size re-estimation, early stopping due to efficacy or futility, and dropping inferior treatment groups. An adaptive design can also involve any combination of these adaptations. Different types of dose increments are discussed including dose escalation and dose reduction, and their combinations. Some statistical methods to control type I error will be presented based on Fisher's combination of independent p-values and hierarchical testing.

68 Issues Related to the Analysis of Longitudinal Data ●▲

Biopharmaceutical Section, Section on Statistics in Epidemiology, Section on Survey Research Methods, WNAR, Biometrics Section

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Likelihood-Based, Mixed-Effects Model Repeated Measures (MMRM) Analysis for Acute Phase Clinical Trial for the Treatment of Major Depressive Disorder

✱Baldeo K. Taneja, EPIX Pharmaceuticals, Inc., 4 Maguire Road, Lexington, MA 02421, *btaneja@epixpharma.com*; Thomas Kinghorn, MDS Biostatistics; John Reinhard, Brain Insights, LLC; Andrew Uprichard, EPIX Pharmaceuticals, Inc.

Key Words: Clinical Trials, Major Depressive Disorder, Longitudinal Data, Missing Data, Dropouts, MMRM Analysis

Statistically valid analyses of longitudinal data can be problematic in the presence of missing values. Even though there are numerous available methods, it is important to realize that no single best method currently exists to deal with this problem. In a recent randomized, double-blind, placebo-controlled 8-week long clinical trial to assess the efficacy and safety of PRX-00023 in patients with major depressive disorder, about one-quarter of the patients withdrew prematurely. On the basis of theoretical and empirical evidence, a likelihood-based mixed-effects model repeated measures (MMRM) analysis was chosen as the primary analysis in the protocol. In this paper, the appropriateness of the design and the related statistical analysis will be highlighted with results.

Longitudinal Data Analysis: Repeated Measures Analysis Versus Random Coefficient Models

✱Radhi Abdulnabi, Ovation Pharmaceuticals, Inc., 4 Parkway North, Deerfield, IL 60015, *rabdulnabi@ovationpharma.com*

Key Words: Longitudinal data, Random coefficient models, Repeated measures, Mixed model

A longitudinal data set is defined as a data set in which the response for each experimental unit is observed on two or more occasions. There are different traditional methods to analyze the longitudinal data including: repeated measures ANOVA and mixed effects models. Random Coefficient Models (RCM) are anther way to analyze the longitudinal data by using the time as continues variable and estimating the slop, intercept, and the regression line of each subject. Clinical trials data with high dropout rate were used to compare the results of RCM versus the traditional methods and to evaluate the missing data effect. Using the time in the model as fixed effect versus random effect or as class versus continuous variable were examined. The final results show that the RCM (using the time as random continues variable) have greater power and better model fitting.

Mixture Model for Individual and Combined Data Using Estimating Equations

✱Pali Sen, University of North Florida, 1 UNF Drive, Dept of Math & Stat, Jacksonville, FL 32224, *psen@unf.edu*; Jacy Crosby, University of North Florida

Key Words: Drug interactions, Stochastic Differential equations, Isomers

In this paper we consider a new analytical framework that is a combination of the individual and combined data analyses, based on an estimating equation approach. The proposed analyses utilize a stochastic model for the two drug combinations and derive the mean and the variance terms based on Ito's calculus. The proposed estimation methods are used to estimate model parameters from individual and combined data provides the ground for the model free tests. The strength of the fit of the model to the data is examined by the statistical measures and the graphical method. Simulation studies were performed to show the strengths of the proposed approach in the estimating the model parameters. A synergy test of the model fitted by the individual subjects confirmed that the combination of the drugs under study is synergic in nature.

A Comparison of Methods for Pretest/Post-Test Trials with Small Samples

✱Xiao Sun, Merck & Co., Inc., UG1CD-44, 351 N. Sumneytown Pike, North Wales, PA 19454, *Xiao_Sun@merck.com*; Robin Mogg, Merck Research Laboratories; Devan V. Mehrotra, Merck Research Laboratories

Key Words: pretest-posttest trials, ANCOVA, GEE, RAVE, non-parametric

The pretest/post-test setting is commonplace in clinical trials. One typical trial objective is to detect difference in the response between two treatment groups. The response variable is measured at or before randomization (pretest) and at prespecified follow-up times (posttest). In search of the most efficient way to analyze this type of data, we compared various common analysis methods typical in Phase I HIV vaccine trials and concluded that the ANCOVA method is generally best with valid assumptions. However, due to small sample sizes and outlying observations, the normality assumption is often suspected. As such, we further looked into semiparametric (GEE), robust parametrics (RAVE) and nonparametric (rank based) methods. Simulations were used to compare the different methods under various conditions, and practical recommendations for statisticians will be highlighted.

MMRM Analyses with and Without Titration Visits

✱Phillip Dinh, U.S. Food and Drug Administration, 10903 New Hampshire Ave, Silver Spring, MD 20993-0002, *phillip.dinh@fda.hhs.gov*; Peiling Yang, U.S. Food and Drug Administration

Key Words: Missing data, MMRM, Repeated measures, Titration

In longitudinal clinical studies, after randomization at baseline, subjects are followed for a period of time for development of symptoms. A mixed model for repeated measures (MMRM) can be used to analyze such data. To accommodate safety and tolerability, in some studies, the treatment has to be titrated to the optimal dose. Then, subjects will stay on their optimal dose until the end of the study. In an MMRM analysis, one can ignore the titration visits because including them could add extra variability unnecessarily. However, when patients drop out during the titration period, ignoring the titration visits would result in missing data in these patients. In this paper, we evaluate the impact of excluding and including titration visits in an MMRM analysis by a simulation study. We evaluate the approaches based on the bias and the coverage accuracy of the confidence interval.

Modeling Phase-Dependent Effects and Volatile Longitudinal Responses via Geometric Brownian Motion Process

✱Li Zhu, University of California, Davis, 4418 Mathematical Sciences Building, Department of Statistics, Davis, CA 95616, *lzhu@wald.ucdavis.edu*; Fushing Hsieh, University of California, Davis; Eric Chi, Amgen, Inc; Juan Li, Amgen, Inc

Key Words: Phase-changes, Profile likelihood, Random effects model, Stochastic differential equation

In this talk we address statistical issues arisen from post marketing evaluations of pharmaceutical products, which focus particularly on disease treatment that gives rise to phase-dependent effects and volatile longitudinal responses. We employ geometric Brownian motion process in this talk to accommodate the two aforementioned longitudinal characteristics. Our model formulation also adapt to firstly finding what individual conditions determine the individual trajectory of disease progression and volatility, and then secondly how to summarize individual information on population level. A two-step modeling approach is proposed, is demonstrated with a data set from a rheumatoid arthritis study in which patients were randomized to three treatment groups and the index measurements disease activity score based on 28 joints (DAS28) were recorded at multiple time points.

Sensitivity Analyses for Data in Presence of Missing Not at Random: A Case Study

✱Yanning Liu, Johnson & Johnson Pharmaceutical R&D, LLC, 1125 trenton-harbourton Road, Titusville, NJ 08628, *yliu@prdus.jnj.com*

Key Words: Last-observation carried forward, Missing not at random, Pattern mixture method, Joint modeling, Longitudinal Analysis

In a one-year noninferiority trial, analysis based on LOCF data resulted a lower bound of -5.84 for the 95% CI of LS means, thus failed to demonstrate the non-inferiority of drug A against drug B within a margin of -5. An overall high dropout rate of 44% was also observed, along with doubling rate of withdrawal due to lack-of-efficacy for A vs. B in the early phase of the trial. Data were showed not to be MCAR by a Chi-squared test and bootstrapping. A series of longitudinal analyses (MMRM, pattern-mixture and joint modeling) are done for sensitivity purposes. Furthermore, a novel method was invented combining advantages from both pattern mixture and joint modeling; this "patterned-joint modeling" largely shorten confidence interval and obtained 95% CI of (-2.00,1.54). Some simulation studies were also done to evaluate empirical type I error of the "pattern-joint modeling" method.

69 Confirmatory Factor Analysis and Principal Components Analysis ●

Social Statistics Section, Section on Nonparametric Statistics, Section on Government Statistics

Sunday, August 3, 4:00 p.m.–5:50 p.m.

Principal Component Analysis for Symbolic Data

✱Jennifer Le-Rademacher, The University of Georgia, 3469 Benthollow Lane, Department of Statistics, Duluth, GA 30096, *jle1@uga.edu*; Lynn Billard, The University of Georgia

Key Words: symbolic data, principal component analysis, interval-valued data, histogram-valued data

In this presentation, we propose a new approach to principal component analysis (PCA) for symbolic data. Symbolic data was first introduced by Edwin Diday in 1987. Unlike classical random variable which takes single values, symbolic variable takes multiple values such as lists, intervals, or distributions. Existing methods of PCA for symbolic data are limited to interval-valued type. However, these methods only account for part of the total variation of interval-valued data. We propose a new method using the so-called symbolic covariance structure to account for the total variation. This approach is further extended to histogram-valued data. We will illustrate with examples.

On Principal Components and Regression: A Statistical Explanation of a Natural Phenomenon

✱Andreas A. Artemiou, The Pennsylvania State University, 325 Thomas Building, University Park, PA 16802, *aaa195@stat.psu.edu*; Bing Li, The Pennsylvania State University

Key Words: Dimension reduction, Orientationally uniform distribution, Principal components, Random covariance matrices, Regression, Stochastic ordering

In this work we give a probabilistic explanation of a phenomenon that is frequently observed but whose reason is not well understood. That is, in a regression setting, the response (Y) is often highly correlated with the leading principal components of the predictor (X) even though there seems no logical reason for this connection. This phenomenon has long been noticed and discussed in the literature, and has received renewed interest recently because of the need for regressing Y on X of very high dimension, often with comparatively few sampling units, in which case it seems natural to regress on the first few principal components of X. This work stems from a discussion of a recent paper by Cook (2007) which, along with other developments, described a historical debate surrounding, and current interest in, this phenomenon.

The Effects of Toeplitz Structure on the Structural Integrity of Confirmatory Factor Analysis

✱Bruce L. Brown, Brigham Young University, 1098 SWKT, Provo, UT 84602, *bruce_brown@byu.edu*; Jeffrey R. Wheeler, Pentara Corp; Tyson T. Phalp, Myriad Genetics; Petr Blahus, Charles University in Prague

Key Words: confirmatory factor analysis, structural integrity, Monte Carlo simulation

Lauritzen et al (2007) demonstrated that factor analysis cannot produce structurally valid results with a reliability less than r = .50. Very few factor analytic studies meet these reliability requirements. Brown et al (2007) demonstrated similar findings with respect to confirmatory factor analysis. Monte Carlo simulation is used to compare the confirmatory factor analysis under two conditions: a clustered configuration of manifest variables and a Toeplitz structure. It is demonstrated that the structural integrity of confirmatory factor analysis is compromised somewhat when the configuration of manifest variables is Toeplitz rather than clustered. Both conceptually and empirically clustered structure is more consistent with the rationale of confirmatory factor analysis than is Toeplitz structure.

Are Husband and Wives' Hostility and Support Distinct Factors?

✱Frederick O. Lorenz, Iowa State University, 2625 North Loop Dr., Suite 500, ISU Research Park, Ames, IA 50011, *folorenz@iastate.edu*; Janet N. Melby, Iowa State University; Rand D. Conger, University of California, Davis

Key Words: confirmatory factor analysis, multitrait-multimethod models, hostility, support, close relationships

Distinct concepts such as hostility and support between wives and husbands or between parents and children are often highly collinear in their measurement, especially when the reports are obtained from self-administered questionnaires. In this paper, confirmatory factor analysis (CFA) is used to estimate multi-trait, multi-method (MTMM) models of self- and spouse- reports of husbands' and wives' hostility and support in a sample of 310 recently married or cohabiting couples. The results suggest that measures of hostility are distinct from measures of support, but that husband and wife reports of hostility, and husband and wife reports of support, are redundant (e.g., in predicting marital outcomes, little of the variance in husbands' reports of their hostility is unique after controlling for wives' reports of husbands' hostility).

Comparison of Methods To Generate Eigenvalues for Parallel Analysis

✱Kellie Keeling, University of Denver, DCB Statistics MSC 8952, 2101 S. University Blvd, Denver, CO 80208, *kkeeling@du.edu*; Robert Pavur, University of North Texas

Key Words: Principal Components Analysis, Parallel Analysis

Parallel Analysis is documented as a preferred method for determining the number of components to retain in a principal components analysis. In order to implement parallel analysis, the eigenvalues from random data are compared to the sample eigenvalues. There have been several published programs that have been proposed to help in the implementation of parallel analysis. This study will compare published programs to see how they compare in eigenvalue generation.

Multigroup Confirmatory Factor Analysis: Parameter Bias and Noninvariant Referents

✱Holmes Finch, Ball State University, Educational Psychology, Ball State University, Muncie, IN 46011, *whfinch@bsu.edu*; Brian F. French, Purdue University

Key Words: Multi-group confirmatory factor analysis, factor invariance, estimation bias, factor analysis

Multigroup confirmatory factor analysis (MCFA) is waxing as a method for examining measurement invariance and specifically, factor invariance. In order to identify the model, MCFA requires certain parameters to be constrained and assumed invariant across groups to act as referent variables. When this invariance assumption is violated, identification of the parameters that actually differ across groups becomes difficult, in part due to the biased estimates resulting from the non-invariant referent. However, there is a dearth of empirical evaluation of the extent to which parameters are biased in the non-invariant referent case. This study examined parameter bias with three identification procedures in the presence of a non-invariant referent. Results suggest that bias is present across all methods, and that bias is particularly evident for factor loadings and inter-factor correlations.

70 Section on Statistical Consulting Roundtable with Coffee (fee event) ▲

Section on Statistical Consulting
Monday, August 4, 7:00 a.m.–8:15 a.m.

Statistical Consulting in the Medical Device Area: Using Secondary Data Sources for Development of Clinical Research Studies

✱Madhu Mazumdar, Weill Cornell Medical College, 411 E. 69th St, KB-323, New York City, NY 10021, *mam2073@med.cornell.edu*

Key Words: Medical Device, Secondary Data Sources, Total Joint Replacement, Clinical Epidemiology, Statistical Consulting, Muscoskeletal disease

The identification and use of large data sources for medical device epidemiologic studies are important. This information helps improve knowledge for public health, regulatory decisions, and progress in clinical research. Limited information about both government and private data sources are available, but they are not sufficient for generating ideas for clinical research, especially for young investigators. A proactive role by the biostatistical consultant in gathering the type of research work already performed and dialogue about what more can be done could help start projects. These projects are typically of better quality as the statistician is involved from the beginning and can control the quality. A scenario in the context of joint replacement (knee, hip, and shoulder) will be discussed. Other examples are welcome.

71 Section on Statistical Education Roundtable with Coffee (fee event)

Section on Statistical Education
Monday, August 4, 7:00 a.m.–8:15 a.m.

Fostering Active Learning in Online and Distance Introductory Statistics Courses

✱Patti B. Collings, Brigham Young University, 231 TMCB, Provo, UT 84602, *collingsp@stat.byu.edu*

Key Words: active learning, GAISE guidelines, Introductory statistics, online course, distance learning course

Are you involved with an online or distance learning introductory statistics course? Have you wondered how to foster active learning in your course? (Fostering active learning is one of the recommendations of GAISE.) Maybe you are already successful in doing this and would like to share your ideas. Maybe you'd like some ideas. Let's get together, brainstorm, share ideas, and talk about the challenges associated with implementing active learning in our online and distance learning courses.

72 Section on Statistics and the Environment Roundtable with Coffee (fee event)

Section on Statistics and the Environment
Monday, August 4, 7:00 a.m.–8:15 a.m.

Get Involved in the Research on Environmental Statistics Going on in RTP

✱Montserrat Fuentes, North Carolina State University, Statistics Department, Raleigh, NC , *fuentes@stat.ncsu.edu*

Key Words: climate change, weather, air pollution, environmental statistics, spatial statistics, computational statistics

Recent scientific evidence seems to indicate major and widespread climate changes have occurred. Are those changes significant? What is the uncertainty associated to the reported numbers? What is the impact of climate change on air pollution? EPA is proposing to strengthen air quality standards for ground-level ozone. What is the role of statisticians in the research that motivates these important policy decisions? In this roundtable, we will identify important areas of research on environmental statistics, the limitations of current statistical methods, computational challenges and how to overcome them, and, most importantly, how you can get involved in this type of research. Join us and find out how statisticians get involved in the most relevant problems in weather, climate, and air pollution—and how RTP is one of the leading centers for this type of research.

73 Section on Statistics in Epidemiology Roundtable with Coffee (fee event)

Section on Statistics in Epidemiology
Monday, August 4, 7:00 a.m.–8:15 a.m.

Bayesian Methods for High-Dimensional Data

✱David Dunson, National Institute of Environmental Health Science, 111 T. W. Alexander Dr., Research Triangle Park, NC 27709, *dunson1@niehs.nih.gov*

Key Words: Bayesian, Regularization, Variable selection, Hierarchical model, Shrinkage, MCMC

In epidemiologic studies, data are often collected for a large number of predictors and health outcomes. Epidemiologists often focus on simple analyses based on selecting a few predictors and outcomes from among the many collected in a study. Such a strategy can produce misleading results and result in an inflated type I error rate. This roundtable will discuss Bayesian methods for analysis of high-dimensional data from epidemiologic studies, focusing on flexible shrinkage and variable selection methods, and practical issues that arise in applying such methods in epidemiologic studies. In addition, we may discuss structural equation methods. Studies collecting SNP data, with an interest in gene x environment interaction, provide one application, though high-dimensional data often arise in epidemiologic studies without a genetic component.

74 Section on Health Policy Statistics Roundtable with Coffee (fee event)

Section on Health Policy Statistics
Monday, August 4, 7:00 a.m.–8:15 a.m.

Multiplicity Issues Related to Composite Endpoint of a Pro Instrument

✱Mahboob Sobhan, U.S. Food and Drug Administration, 10903 New Hampshire Av, Silver Spring, MD 20993, *mahboob.sobhan@fda.hhs.gov*

Key Words: Composite endpoint, Multiplicity

Statistical issues related to testing of a single prespecified component of a composite endpoint will be discussed. Specifically, what statistical strategies are best to make a claim when a composite is also a coprimary endpoint.

75 Section on Survey Research Methods Roundtable with Coffee (fee event)

Section on Survey Research Methods
Monday, August 4, 7:00 a.m.–8:15 a.m.

Surveying Cell Phone Numbers in the United States

✱Paul J. Lavrakas, Consultant, 382 Janes Lane, Stamford, CT 06903, *pjlavrak@optonline.net*

Key Words: cell phone, RDD, telephone survey

How to handle cell phone numbers has become a major concern in telephone surveying in the US. This roundatble will discuss key issues in coverage, sampling, nonresponse, measurement, and weighting that affect U.S. telephone surveys that reach cell phone numbers, especially surveys meant to sample the general public with an RDD frame. We also will discuss unique legal and ethical issues that come into play when reaching U.S. cell phone numbers, compared to what happens when dialing U.S. landline numbers. In addition, operational considerations (including cost implications) will be included in the discussion. The information presented at the roundtable will include updates from new research and thinking since the publication of the December 2007 special issue of Public Opinion Quarterly on “Cell Phone Numbers and Telephone Surveying in the US.”

76 Section on Quality and Productivity Roundtable with Coffee (fee event)

Section on Quality and Productivity
Monday, August 4, 7:00 a.m.–8:15 a.m.

Statistics Groups in Industry: Rise, Decline, and Improvement

✱Gerald Hahn, GE Corporate Research and Development (retired), Applied Statistics, 1404 Orlyn Drive, Schenectady, NY 12309, *gerryhahn@yahoo.com*

Key Words: central statistics groups, industrial statistics, consulting, metrics for success, communications

Various companies (and other organizations) have formed statistics groups over the years to provide a centralized resource. This roundtable will consider such groups and address the following inter-related questions: What should be the role of a central statistics group? What are the advantages and limitations of such groups, versus individually dispersed statisticians, to the organization and the individuals involved? How should such a group be funded? Why have such groups declined in U.S. industry over recent years? What are the lessons learned from this decline, and what should/can we do about it? What are key pointers to success for managers and members of such groups? What are pitfalls? How can “success” be measured and communicated effectively to management?

77 Introductory Overview Lecture: Missing and Coarse Data

ASA, ENAR, IMS, SSC, WNAR, Biopharmaceutical Section, Section on Survey Research Methods, Section on Teaching Statistics in the Health Sciences, Section on Statistics in Epidemiology
Monday, August 4, 8:30 a.m.–10:20 a.m.

An Introduction to Multiple Imputation

✱Roderick J. Little, The University of Michigan, Department of Biostatistics, School of Public Health, 1420 Washington Hgts, M4208 SPH II, Ann Arbor, MI 48109, *rlittle@umich.edu*

Imputation is a common pragmatic approach for handling missing data—missing values are replaced by estimates and analyses are conducted on the filled-in data set. The obvious drawback of imputation is that it is “cheating,” since it “makes up” data. As a result, inferences from the filled-in data overstate the amount of information, yielding confidence intervals that are too narrow, and tests that do not achieve nominal levels. Multiple imputation addresses this concern by creating multiple data sets with different sets of plausible values imputed. The basic idea and underlying theory of multiple imputation will be reviewed, including multiple imputation combining rules, imputation models, controversies about the method, available software, and pros and cons relative to other missing data approaches.

Analyzing Coarse Data

✱Daniel F. Heitjan, University of Pennsylvania, Biostatistics & Epidemiology, 622 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, *dheitjan@mail.med.upenn.edu*

Key Words: Heaped data, ignorability, incomplete data

Often we can think of our data as an imperfectly observed version of a hypothetical complete data set in which some or all observations are known only imprecisely. When the degree of imprecision is so great that standard continuous-data models are no longer reasonable approximations, we say that the data are coarse. Some prominent examples of coarse data include censored data (e.g., survival times in clinical trials), grouped data (as incomes), and heaped data (data with multiple degrees of coarseness, such as reported ages or daily cigarette consumption). I will present a general analysis strategy for

coarse data, discuss ignorability conditions, and illustrate the methodology with one or two live examples.

78 Current Issues in Molecular Epidemiology: Heterogeneity and High-Dimensionality ●

Section on Statistics in Epidemiology, WNAR, Biometrics Section

Monday, August 4, 8:30 a.m.–10:20 a.m.

Methods for Incorporating Biological Knowledge into Analysis of Genome-Wide Association Studies

✱Hongzhe Li, University of Pennsylvania, Department of Biostatistics and Epidemiology, 920 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, *hongzhe@mail.med.upenn.edu*

Key Words: group additive regression, pathways and networks, hidden Markov random fields, protein-protein interaction

Genome-wide association (GWA) studies have become increasingly popular as a powerful tool for the identification of the disease-causing germline genetic variants. The commonly used analytic method has been mainly single SNP or SNP-SNP pair analyses, coupled with statistical strategies for controlling multiple comparisons. However, this simple approach can lead to both false positive and false negative identification of the relevant SNPs due to both high-dimensionality of the data and the complex genetic architecture of complex diseases. One solution to this problem is to integrate other biologically relevant information into the analysis of such GWA studies. I will present several statistical methods for incorporating the SNP/gene annotations and the pathways information into analysis of these data and demonstrate the methods using a case-control GWA study of neuroblastoma.

Probability of Detecting Disease-Associated SNPs in Case-Control, Genome-Wide Association Studies

✱Mitchell H. Gail, National Cancer Institute, 6120 Executive Blvd., Room 8032, Bethesda, MD 20892-7244, *gailm@mail.nih.gov*; Ruth Pfeiffer, National Cancer Institute; William Wheeler, Information Management Services; David Pee, Information Management Services

Key Words: genetic association study, genome-wide association study, detection probability, single nucleotide polymorphism, ranking and selection

Some case-control genome-wide association studies select promising single nucleotide polymorphisms (SNPs) by ranking corresponding p-values. We define the detection probability (DP) for a specific disease-associated SNP as the probability that the SNP will have one of the T smallest p-values for trend tests of association. The proportion positive (PP) is the fraction of selected SNPs that are true disease-associated SNPs. DP increases with genetic effect size and case-control sample size, and decreases with the number of non-disease-associated SNPs. DP increases very slowly with T. For a genetic odds ratio per minor disease allele of 1.2 or less, even a study with 1000 cases and 1000 controls requires T to be impractically large to achieve an acceptable DP and PP. These ideas are also used to minimize the total cost of a research program that follows up on the selected SNPs.

Shrinkage Estimators for Robust and Efficient Inference in Haplotype-Based Case-Control Studies

✱Yi-Hau Chen, Institute of Statistical Science, Academia Sinica, No 128 Sec 2 Academia Road, Taipei, 11529 Taiwan, *yhchen@stat.sinica.edu.tw*; Nilanjan Chatterjee, National Cancer Institute; Raymond Carroll, Texas A&M University

Key Words: bias-efficiency trade-off, empirical Bayes, gene-environment interaction, penalized likelihood

We develop methods for haplotype-based analysis of case-control studies which gain efficiency by exploiting model assumptions of HWE and G-E independence and yet are resistant to bias when those assumptions are violated. The idea involves shrinkage of a “model-free” estimator that is robust to HWE and G-E independence towards a “model-based” estimator that directly exploits those assumptions. A “model-free” estimator is first proposed which is closely related to the “prospective” estimator. Considering the “retrospective” estimator in Spinka et al. (2005) as the “model-based estimator, we then propose shrinkage estimators based respectively on the empirical Bayes and the penalized likelihood approaches. Variance estimation for these shrinkage estimators is proposed. Bias vs. efficiency trade-off for the shrinkage estimators is studied via simulations, and a real application is provided.

Assessing Gene-Environment Interaction in a Genomewide Association Study

✱James Gauderman, University of Southern California, 1540 Alcazar St, Suite 220, Los Angeles, CA 90089, *jimg@usc.edu*; Cassandra Murcray, University of Southern California; Juan Lewinger, University of Southern California; Heather Volk, University of Southern California; David Conti, University of Southern California; Dalin Li, University of Southern California

Key Words: interaction, environment, association, gene

Genomewide association (GWA) studies are designed to test several hundred thousand single nucleotide polymorphisms (SNPs) for association with a trait. Most testing procedures focus solely on detecting the direct effect of each SNP on the trait. However, for complex human traits, it is commonly believed that gene-environment (GxE) interactions play an important role in etiology. We will review commonly used approaches for detecting GxE interactions. We will also present novel approaches, designed specifically for the GWA context, to efficiently identify genes involved in a GxE interaction. Applications to an ongoing GWA study of asthma and lung function in a large cohort of children will be provided.

79 Financial Econometrics

Business and Economics Statistics Section

Monday, August 4, 8:30 a.m.–10:20 a.m.

Inference for Lévy-Driven, Continuous-Time ARMA Processes

✱Richard A. Davis, Columbia University, Department of Statistics, 1255 Amsterdam Ave, MC 4690, New York, NY 10027, *rdavis@stat.columbia.edu*; Peter J. Brockwell, Colorado State University; Yu Yang, Colorado State University

Continuous-time ARMA (CARMA) processes with non-negative kernel and driven by nondecreasing Lévy processes constitute a very general class of stationary, non-negative continuous-time processes. In financial econometrics, for example, they have been used to model stochastic volatility (e.g., Barndorff-Nielsen and Shephard (2001) and Todorov and Tauchen (2006)). In this paper, we develop a highly efficient method of estimation for the coefficients of such models, taking advantage of the non-negativity of

the driving process. We also show how to reconstruct the background driving Lévy process from a continuously observed realization of the CARMA process and use this result to estimate the increments of the Lévy process itself when closely spaced observations are available.

Microstructure Noise, Integrated Volatility, and Rounding Error

✱Mathieu Rosenbaum, University Paris-Est and CREST-ENSAE, Timbre J 120, 3 rue Pierre Larousse, Malakoff, 92245 France, *rosenbaum@ensae.fr*

Key Words: Microstructure noise, Diffusion models, Integrated volatility, High frequency data, Round-off error, Variation methods.

We consider a microstructure model for a financial asset, allowing for prices discreteness and for a diffusive behavior at large sampling scale. This model consists in the observation at the high frequency n, with rounding error a(n), of a diffusion on a finite time interval. From this sample, we give estimators of the absolute and relative integrated volatilities of the asset. Our method is based on the use of variational properties of the process in a wavelet setting. We prove the accuracy of our estimation procedures is max(a(n), n^{-1/2}). Using compensated estimators, limit theorems are given in the case of a homogeneous diffusion coefficient.

A Levy-Driven, Continuous-Time GARCH Process

✱Alexander Lindner, TU Braunschweig, Institute for Mathematical Stochastics, Pockelstrasse 14, Braunschweig, International 38106 Germany, *a.lindner@tu-bs.de*

Key Words: COGARCH, GARCH, Levy process

A continuous time GARCH process which is driven by a Levy process is introduced. It is shown that this process shares many features with the discrete time GARCH process. In particular, the stationary distribution has heavy tails. Extensions of this process are also discussed. We then turn attention to some first estimation methods for this process, with particular emphasis on a generalized method of moment estimator. Finally, we also report on how the continuous time GARCH process approximates discrete time GARCH processes when sampled at discrete times. The talk is based on joint work with Stephan Haug (TU Munich), Claudia Klueppelberg (TU Munich) and Ross Maller (Australian National University).

80 In Remembrance of Paul Minton: Statistician, Educator, and Advocate for Statistics ▲

Memorial, ENAR, Committee on ASA Archives and Historical Materials

Monday, August 4, 8:30 a.m.–10:20 a.m.

Paul Dixon Minton: LSD for Statisticians from a Southern Gentleman

✱James M. Davenport, Virginia Commonwealth University, Department of Stat. Sciences & Ops. Res., P. O. Box 843083, Richmond, VA 23284-3083, *jdavenpo@vcu.edu*

Key Words: Educator, Service, Advocate, Leader

Paul Dixon Minton joined Virginia Commonwealth University (VCU) during the summer of 1972 as the Dean of the School of Humanities and Sciences. It was a time of change and transition as VCU had been created in 1968 by and Act of the Virginia Legislature. Dr. Minton provided the leadership that was needed and admirably served the university as Dean until 1978, then as Director of the Institute of Statistics before retiring in 1988. A dominate theme in his career is his dedication to serving all of his constituents. He was a recipient of the prestigious American Statistical Association's Founders Award (1991), and the namesake of the Paul D. Minton Service Award given by the Southern Regional Council on Statistics (SRCOS). An anecdotal remembrance of several events in Dr. Minton's years at VCU and his years of participation with SRCOS will honor this kind gentleman who loved his puns.

The 'Compleat' Statistician: Paul Minton

✱Fritz Scheuren, The University of Chicago, NORC, 1402 Ruffner Rd, Alexandria, VA 22302, *scheuren@aol.com*

This paper is about a statistician, Paul Minton, who was "compleat" in a way that the old-fashioned spelling implies. My memories of him will be the main source here of these remarks, as I assume will be the case of the others who speak. There will be some repetition in this approach. But that is all to the good—for emphasis, if nothing else. We have a man who played many professional roles with distinction: teacher, scholar, and mentor. In this talk, naturally, I will address these. Hardest to talk about and most elusive is the way he moved seamlessly from one role to another. You always felt a sense of balance as between his personal and professional live(s). How he did this and how he tried to "example it" are my biggest memories. His gift eludes me yet. Maybe together today in this session we will arrive at an understanding that will allow us to continue his legacy.

Development and Growth of Statistical Science at Southern Methodist University

✱William R. Schucany, Southern Methodist University, PO Box 750332, Dept of Statistical Science, Dallas, TX 75275-0332, *schucany@smu.edu*

Key Words: education, student, recruiting, faculty, MS, PhD

We at SMU are the beneficiaries of a department that was founded by Paul Minton. He led the development of two graduate degree programs after recruiting an entire faculty. He fostered other professional activities such as a statistical consulting lab and the North Texas Chapter of the ASA. He was a teacher, who had his eye out for talented statisticians. He was a talented administrator. Paul was always a great leader in service to and service by our profession.

81 Disparate Information Fusion ●

Section on Statisticians in Defense and National Security, Section on Physical and Engineering Sciences

Monday, August 4, 8:30 a.m.–10:20 a.m.

On the Exploitation of Multiple Disparate Dissimilarities

✱Carey E. Priebe, Johns Hopkins University, Baltimore, MD 21218-2682, *cep@jhu.edu*; Youngser Park, Johns Hopkins University; Zhiliang Ma, Johns Hopkins University; Adam Cardinal-Stakenas, Johns Hopkins University

Key Words: interpoint comparisons, interpoint distances, pattern recognition, classification, shape analysis, LDDMM

We investigate aspects of statistical inference and statistical pattern recognition associated with observing multiple disparate dissimilarities between entities rather than observing feature vectors associated with the individual entities themselves. In particular, we consider methods for comparing and

combining dissimilarity matrices obtained by using disparate dissimilarity measures when the exploitation task of interest is classification. We demonstrate our ideas in the context of disease investigation using brain shape comparison data.

Embedding Methods for Disparate Data

✱Michael W. Trosset, Indiana University, P.O. Box 6424, Bloomington, IN 47407, *mtrosset@indiana.edu*; Brent S. Castle, Indiana University

Key Words: multidimensional scaling

When objects (e.g., web pages) comprise disparate elements (e.g., text and pictures), it may not be clear how to construct a suitable representation of the objects for subsequent analysis. One possibility is to measure pairwise dissimilarities with respect to each disparate element, then extract a single Euclidean representation of the objects by three-way (individual differences) nonmetric multidimensional scaling. We discuss various technical difficulties with this approach, propose ways of overcoming those difficulties, and apply our methods to actual data.

Some Strategies for the Fusion of Imagery and Text

✱Jeffrey L. Solka, Naval Surface Warfare Center, Warfare Center Dahlgren Division, Code Q21, 17320 Dahlgren Rd., Dahlgren, VA 22448-5100, *jeffrey.solka@navy.mil*; Nick Tucey, Naval Surface Warfare Center, Warfare Center Dahlgren Division; Avory Bryant, Naval Surface Warfare Center, Warfare Center Dahlgren Division; Ted Clark, Naval Surface Warfare Center, Warfare Center Dahlgren Division

Key Words: text, image, fusion, projection, clustering, visualization

This talk extends some of our original work in the area of fusion of disparate text corpora into the area of image text fusion. We will discuss our methodologies for text/image simultaneous collection along with a number of ideas including utilization of text features as surrogates for image features, techniques for the simultaneous projection of text and imagery data, and other techniques that seek to use both text and image features as part of clustering/pattern recognition schemes. The developed methodologies will be illustrated on a set of image/text news data.

82 Robust Methods in Small-Area Estimation ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section, Section on Health Policy Statistics
Monday, August 4, 8:30 a.m.–10:20 a.m.

Examining Sensitivity of Small-Area Inferences to Uncertainty About Sampling Error Variances

✱William R. Bell, U.S. Census Bureau, Room 5K142A, SRD, 4600 Silver Hill Road, Washington, DC 20233, *William.R.Bell@census.gov*

Key Words: sampling error model, Fay-Herriot model, Bayesian inference, small area estimation

Small-area estimation based on area level models typically assumes that sampling error variances for the direct survey small area estimates are known. In practice we use estimates of the sampling error variances, and these can contain substantial error. To account for this error in small area estimation we present a Bayesian approach that combines a sampling error model developed from the direct sampling variance estimates with a Fay-Herriott type model for the direct survey point estimates. We apply the approach to models for state age group poverty rates, and their sampling error variances, from the Census Bureau's Small Area Income and Poverty Estimates program. We thus examine how uncertainty about the sampling error variances affects the model-based inferences about the true poverty rates.

Robust Estimation of Monthly Employment Growth Rates

✱Julie Gershunskaya, Bureau of Labor Statistics, 2 Massachusetts Avenue, NE, Washington, DC 20212, *gershunskaya.julie@bls.gov*; Partha Lahiri, University of Maryland

Key Words: small area estimation, robust estimation, influential observations

Each month, the Bureau of Labor Statistics publishes estimates of employment for industrial supersectors at the metropolitan statistical area (MSA) level. The survey-weighted ratio estimator that is used to produce estimates for larger domains is less reliable for MSA level of estimation due to small sample sizes. We also note that the effect of data from a few establishments, which are influential in terms of unusual employment numbers or sampling weights or both, is more prominent for the small areas. In this paper, we develop a model-based method that is less variable and less sensitive to influential establishments when compared to the estimator currently used by the BLS.

Robust Small-Area Estimation Under Unit-Level Models

✱J. N. K. Rao, Carleton University, School of Mathematics and Statistics, Ottawa, ON K2G 4H8 Canada, *jrao34@rogers.com*

Key Words: unit level models, resistant methods, outliers, parametric bootstrap, mean squared error, random effects

Empirical best linear unbiased prediction (EBLUP) estimators of small area means have been obtained under unit level nested error regression models. But EBLUP estimators can be highly influenced by the presence of outliers in the data. We propose a resistant method for small area estimation which is useful for downweighting any influential observations in the data when estimating small area means. A parametric bootstrap method is used to estimate the mean squared error (MSE). A simulation study is conducted to study the efficiency of the proposed robust estimators relative to EBLUP estimators and the relative bias of the bootstrap MSE estimators in the presence of outliers. The proposed robust method is also applied to some real data reported in the published literature.

83 High-Dimensional Data Assimilation and the Ensemble Kalman Filter

Section on Bayesian Statistical Science, Section on Statisticians in Defense and National Security
Monday, August 4, 8:30 a.m.–10:20 a.m.

Dynamic Tomography of the Solar Corona with the Localized Ensemble Kalman Filter

✱Richard A. Frazin, The University of Michigan, 2455 Hayward, rm. 1411c, Ann Arbor, MI 48109, *rfrazin@umich.edu*; Yuguo Chen, University of Illinois at Urbana-Champaign; Mark Butala, University of Illinois; Farzad Kamalabadi, University of Illinois

Key Words: Kalman filter, particle filter, data assimilation

The corona is the outermost layer of the Sun's atmosphere, and it is critical to understanding both space weather as well as underlying issues in space plasma physics. Satellites routinely make images of the corona and these images are 2D projections of the 3D corona. The corona is rotating, thus providing a diversity of view-angles, but it is also involving in time. Assimilating the projection data is a challenging problem because one must treat the line integral operators, which are not as easy to localize. These challenges will be discussed and recent progress will be reviewed.

Using Small Ensembles in High Dimensions: Hierarchical Bayesian Approaches to Adaptive Ensemble Filters

✱Jeffrey Anderson, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, *jla@ucar.edu*

Key Words: Ensemble filters, Data assimilation, Hierarchical Bayesian, Adaptive algorithms

Ensemble filters with sample sizes as small as 20 are routinely applied to geophysical models with millions of state vector elements. Sampling error leads to state estimates with too little variance and erroneous correlations between observations and state. Hierarchical Bayesian algorithms that automatically detect and correct for sampling errors as well as some errors in the prediction model are presented. An algorithm to correct variance uses a deterministic filter in concert with an ensemble filter. An algorithm to correct correlations uses an ensemble of ensemble filters. Results will be shown for numerical weather prediction applications.

Estimation of High-Dimensional Prior and Posterior Covariance Matrices in Kalman Filter Variants

✱Reinhard Furrer, Colorado School of Mines, Mathematical and Computer Sciences Dept., 1500 Illinois Street, Golden, CO 80401, *rfurrer@mines.edu*; Thomas Bengtsson, Bell Labs, Alcatel-Lucent

Key Words: Ensemble Kalman filter, Square-root filter, Matrix expansions, Tapering, Shrinking, Covariance boosting

This work studies the effects of sampling variability in Monte Carlo-based methods to estimate very high-dimensional systems. As the employed sample sizes are typically several orders of magnitude smaller than the system dimension, ordinary sampling techniques inevitably induce considerable variability into the state estimate, primarily through prior and posterior sample covariance matrices. We quantify this variability with mean squared error measures and expressions of the error measures are derived under weak assumptions. To reduce necessary ensemble size requirements and to address rank-deficient sample covariances, tapering of the prior sample covariance with a compactly supported positive definite function is demonstrated to be a simple, computationally feasible, and very efficient technique. Rules for obtaining optimal taper lengths are given.

Ensemble Sampling

✱Nicholas Polson, The University of Chicago, 5807 Woodlawn Ave, Chicago, IL 60637, *ngp@chicagogsb.edu*; Hedibert Lopes, The University of Chicago

Key Words: Bayesian, Ensemble, Stochastic Volatility, Filtering, Simulation, Posterior

This paper provide an ensemble sampling approach to Bayesian inference. The method provides approximate samples from complicated posterior distributions for a wide class of models. It does not rely on MCMC or resampling methods and naturally extends to multiple dimensions. We illustrate our methodology for a number of applications including multivariate stochastic volatility models.

84 Ridge Regression and Related Topics ▲

SSC, Section on Physical and Engineering Sciences, IMS

Monday, August 4, 8:30 a.m.–10:20 a.m.

A New Class of Generalized Bayes Minimax Ridge Regression Estimators

✱William E. Strawderman, Rutgers, The State University of New Jersey, 110 Frelinghuysen Rd, Piscataway, NJ 08854, *straw@stat.rutgers.edu*

Key Words: Ridge Regression, Minimax, Generaized Bayes, Squared Error Loss

We consider estimation of the regression parameters in a general linear model, under scale invariant quadratic loss, when the error vector has a spherically symmetric distribution. We extend results of Strawderman (1978) and Casella (1980, 1985) by finding adaptive minimax estimators with greater numerical stability (smaller condition number) than the usual least squares estimator. Under Normality assumptions they are also Generalized Bayes, and under additional assumptions they remain generalized Bayes for essentially all spherically symmetric error distributions. One such subclass has a particularly simple form. (This work is joint with Yuzo Maruyama, University of Tokyo.)

Developing Ridge Parameters for SUR Model

✱Ghazi Shukur, Jonkoping University, Department of Economics and Statistics, P.O. Box 1026, SE- 551 11 Jonkoping, International SE- 551 11 Sweden, *ghazi.shukur@jibs.hj.se*

Key Words: Multicollinearity, SUR ridge regression, Monte Carlo simulations, biased estimators, Generalized least squares

In this paper, a number of procedures have been proposed for developing new biased estimators of the seemingly unrelated regression (SUR) parameters, when the explanatory variables are affected by multicollinearity. Several ridge parameters are proposed and then compared in terms of the trace mean squared error (TMSE) and (PR) criteria. The PR criterion is the proportion of replication (out of 1,000) for which the SUR version of the generalized least squares, (SGLS) estimator has a smaller TMSE than others. The study has been made using Monte Carlo simulations where the number of equations in the system, the number of observations, the correlation among equations and the correlation between explanatory variables have been varied. For large samples and when the colinearity is not high, the unbiased SUR estimator, (SGLS), performed better than the other ridge parameters.

A Class of Unrestricted Estimators

✱M. Revan Özkale, Çukurova University, Faculty Science and Letters, Department of Statistics, Adana, 01330 Turkey, *mrevan@cu.edu.tr*

Key Words: Multicollinearity, Biased estimators, Mean square error

The purpose of this paper is to combine several regression estimators (ordinary least squares (OLS), ridge, contraction, principal components regression (PCR), Liu, r-k and r-d class estimators) into a single estimator. The conditions for the superiority of this new estimator over the PCR, the r-k class, the r-d class and fl(k,d) (proposed by ÷zkale and Kaçiranlar (2007)) estimators are derived by the scalar mean square error criterion and the estimators of the biasing parameters for this new estimator are given. Also, a numerical example based on Hald data is used to illustrate the results.

85 Methodological Advances in Testing and Estimation of Gene Expression Differences ●

ENAR, Section on Nonparametric Statistics, Biopharmaceutical Section, Section on Statistics in Epidemiology, WNAR, IMS, Biometrics Section
Monday, August 4, 8:30 a.m.–10:20 a.m.

Estimating FDR Using Nonparametric Deconvolution

✱Mark van de Wiel, Vu University, De Boelelaan 1081a, Amsterdam, 1081 HV The Netherlands, *mark.vdwiel@vumc.nl*; Kyung In Kim, Eindhoven University of Technology

Key Words: Deconvolution, Gene expression, FDR, Effect size

Given a set of microarray data, the problem is to detect differentially expressed genes, using a false discovery rate (FDR) criterion. As opposed to common procedures, we do not base the selection on statistical significance only, but also on effect size. Therefore, we select only those genes that are significantly more differentially expressed than some f-fold. This corresponds to use of an interval null domain for the effect size. Based on a simple error model, we discuss a naive estimator for the FDR, interpreted as the probability that the parameter of interest lies in the null-domain (e.g., $\mu < \log_2 2 = 1$) given that the test statistic exceeds a threshold. We improve the naive estimator by using deconvolution. That is, the density of the parameter of interest is recovered from the data. We study performance of the methods using simulations and real data.

Exploring the Information in P-Values for the Analysis and Planning of Multiple-Test Experiments

David Ruppert, Cornell University; ✱Dan Nettleton, Iowa State University, 111 Snedecor Hall, Department of Statistics, Ames, IA 50011-1210, *dnett@iastate.edu*; J.T. Gene Hwang, Cornell University

Key Words: microarray, spline, power and sample size, false discovery rate, expected discovery rate, simultaneous testing

A new methodology is proposed for estimating the proportion of true null hypotheses in a large collection of tests. Each test concerns a single parameter delta whose value is specified by the null hypothesis. We combine a parametric model for the conditional cumulative distribution function of the p-value given delta with a nonparametric spline model for the density of delta under the alternative hypothesis. We estimate the proportion of true null hypotheses and the density of delta when the null is false by penalized and constrained least squares. The estimator is computed efficiently using quadratic programming. We discuss the use of our estimate of the density of delta in sample size calculations for future microarray experiments. We compare our estimator to leading competitors through simulation and illustrate our method using microarray data examples.

An Approximate Empirical Bayes Model Selection Approach to Microarray Data Analysis

✱Harrison Zhou, Yale University, New Haven, CT , *huibin.zhou@yale.edu*; J.T. Gene Hwang, Cornell University; Dan Nettleton, Iowa State University

Key Words: model selection, microarray data analysis, sparse inference

Recent approach to model selection is to do variable selection and then estimate the coefficients of the selected variables. Some statistical procedures have been proposed to achieve these two goals simultaneously. In the regression context, it would be theoretically desirable if the proposed procedure dominates the ordinary least squared estimate, namely the natural procedure without variable selection. This property is called minimaxity. In this paper, we show that many well-known procedures fail to be minimax. We construct such minimax estimator which does variable selection as well. On the practical side, the estimators we construct perform as well as other well-known procedures even in the very sparse situations where the coefficients of variables are mostly zero. The procedure is easy to implement and computationally less intensive.

Multiple Testing on the Directed Acyclic Graph of Gene Ontology

✱Jelle Goeman, Leiden University Medical Center, Medical Statistics (S5-P), P.O. Box 9600, Leiden, International 2300 RC The Netherlands, *j.j.goeman@lumc.nl*; Ulrich Mansmann, University of Munich

Methods that test for differential expression of Gene Ontology (GO) terms have so far not been able to make use of the graph structure of Gene Ontology when adjusting for multiple testing. We propose a multiple testing method, called the focus level procedure, that preserves the graph structure of Gene Ontology when testing for association of the expression profiles of GO terms with a response variable. The procedure is constructed as a combination of a Closed Testing procedure with Holm's method. It allows a user to choose a "focus level" in the GO graph, which reflects the level of specificity of GO in which the user is most interested and also determines the level in the GO graph at which the procedure has most power. The procedure strongly keeps the family-wise error rate without any additional assumptions on the joint distribution of the test statistics used.

86 Political Science, Statistical Science, and Graphics ●▲

Section on Statistical Graphics, Section on Survey Research Methods, Social Statistics Section
Monday, August 4, 8:30 a.m.–10:20 a.m.

A Picture Is Worth a Billion Words: Visualizing Mega-Parameter Models from Giga-Scale Textual Data

✱Burt Monroe, The Pennsylvania State University, Dept of Political Science-203 Pond Lab, University Park, PA 16802, *burtmonroe@psu.edu*

This talk discusses model visualization methods that have been developed for a project analyzing records of political speech in legislatures. These records contain billions of words (more than Wikipedia) uttered over timescales from minutes to centuries. Our statistical models—dynamic Bayesian mixture models to study topic attention, dynamic feature selection models to study political content, and dynamic scaling and item response models to study political positioning—produce millions of meaningful parameters. Model visualization is central to both interpretation and communication of these results.

Ceci N'Est Pas Une Carte

✱Michael D. Ward, University of Washington, Political Science, Gowen Hall, Seattle, WA 98195, *mdw@u.washington.edu*

Key Words: cartogram, social science

Use of cartograms and other amorphologic representations of data has been important in studies of the World Wide Web since the last decade. Cartograms have also been around for a long time, but are more frequently seen in contemporary journalism, since the invention of the Newman-Gastner diffusion algorithm. The use of statistics in social science is often to describe patterns that exist on a real geography, but for the most part neither geographic (morphologic) nor representational (amorphologic) graphics have found their way into the standard toolkits of social scientists. I explore the use of amorphologic presentations of a variety of political/geographical data.

Tufte Without Tears

✱Chris Adolph, University of Washington, CSSS, Box 354320, Padelford Hall, Seattle, WA 98195-4320, *cadolph@u.washington.edu*

Key Words: regression, graphics, R package

This talk will present graphical techniques for interpreting and presenting regression and regression-like models. The focus will be on applications in political science; however the results are applicable to a wide audience. Additionally, the visual displays are easily implemented using an R graphics package called "tile," written by the speaker.

87 Statisticians Impacting Policy and Practice in the Real World: Case Studies from Medicine, Fisheries, and Banking ●▲

Committee on Applied Statisticians, Biopharmaceutical Section, Section on Government Statistics, Social Statistics Section, Scientific and Public Affairs Advisory Committee, *CHANCE*

Monday, August 4, 8:30 a.m.–10:20 a.m.

Stochastic Modeling of Oyster Demographics in Support of an Ecological Risk Assessment To Address Management of the Chesapeake Bay Fishery

Mary C. Christman, University of Florida; ✱Thomas Bohrmann, University of Florida, Department of Statistics, *Bohrmann@ufl.edu*

Key Words: hierarchical model, uncertainty, decision-theoretic, Bayesian

Maryland has been considering several alternative management schemes to restore the oyster fishery in the Chesapeake Bay, including introduction of a non-native species and restoration of the existing species. In support of this, we developed a spatially explicit demographic model that can be used to estimate the probability of achieving the stated goal of establishing a naturalized, reproducing, and self-sustaining population of oysters and of increasing biomass levels to 1970 levels within a specified time period. We describe a Bayesian hierarchical model, parameterization of the model, and the incorporation of uncertainty. For uncertainty, we consider the decomposition of uncertainty into two components: natural stochastic variation and uncertainty with respect to parameterization. We use Monte Carlo simulations to assess management alternatives.

Statisticians Influencing Policy and Practice in Banking

✱Leonard D. Roseman, Capital One Services, Inc., Attn: 12073-0300, 15000 Capital One Drive, Richmond, VA 23238, *leonard.roseman@capitalone.com*

Key Words: oral communication, influencing decisions, business, banking

The purpose of statistical analysis is to drive decisions under uncertainty. In banking, as in many business environments, the statistician's role is often to influence decisionmakers. Successful influence requires clear and compelling communication to describe the problem, enumerate solutions, and expose concomitant uncertainties. Statisticians are highly trained in analyzing uncertainty but are rarely trained in the art of persuasive communication. Such additional training is critical to guiding decisionmakers to correct actions and thus critical to making statisticians effective in driving businesses success. I will describe a program that teaches the fundamentals of influential oral communications: brevity, structure, and delivery. We have found this program helps statisticians influence decisionmakers even as it helps them become leaders and decisionmakers themselves.

Bedside Analysis of Cerebral Autoregulation in Very-Low-Birth-Weight Infants

✱D. Keith Williams, University of Arkansas for Medical Sciences, 4301 W Markham slot 781, Little Rock, AR 72205, *williamsdavidk@uams.edu*

Key Words: partial least squares, principal components regression, soft modeling

Advances in medical care have led to increased survival for premature infants. Cerebral autoregulation is an essential physiologic process that maintains constant cerebral blood flow (CBF) to the brain despite alterations in blood pressure (BP). The determination of impaired CBF regulation during the first hours of life prevents brain damage and is a valuable diagnostic tool that optimizes clinical decisions for these infants. Cerebral autoregulation must be determined with a statistical model estimated from an infant's concurrently observed CBF and BP responses over time. Quality predictions are needed to develop an autoregulation parameter. The author presents an autoregulation statistic estimated from continuous biological monitoring system on premature newborns.

88 Applications of Advanced Lifetime Data Analysis ●

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Monday, August 4, 8:30 a.m.–10:20 a.m.

Advanced Reliability Methods for the Optimization of Aircraft Maintenance Process: Cost Minimization

✱I-Li Lu, The Boeing Compnay, P.O. Box 3707 MC 7L-22, Applied Statistics, Phantom Works, Seattle, WA 98124-2207, *i-li.lu@boeing.com*; Ranjan K. Paul, The Boeing Company

Key Words: Gamma Distributions, Extreme Values, Reliability Models, Recurrent Events, Interval Censored Lifetime

In airline maintenance programs, mean cumulative function (MCF) has often been used to estimate the expected number of failures or removals within a given time period to determine the overall cost of aircraft maintenance process. When underlying lifetimes is exponential, these measures provide adequate estimates for the parameters. However, problem arises when lifetimes are no longer exponential. In this talk, we consider fitting

Weibull models to lifetimes extracted from the Boeing Airplane Reliability and Maintainability System database. We use estimated Weibull models and stochastic upper/lower bounds derived from approximation theorems and renewal theory to estimate the expected number of failures for a task or system at a given time. We discuss how these estimates can be applied in cost assessment models implemented in aircraft maintenance programs.

Detection of Nuclear Material in Containers Entering US: A Learning Approach for Analyzing Radiation Portal Data

✱Siddhartha R. Dalal, Rand Corporation, 1776 Main Street, Santa Monica, CA 90401, *sdalal@rand.org*

Key Words: Machine Learning, Illicit Nuclear Material, Radiation, Optimal Decision Rule, Poisson Process

Given the mandate by U.S. Congress that all goods entering US should be inspected for illicit nuclear material, DHS, Department of Homeland Safety, is moving towards one hundred percent inspection of all containers entering US at various ports of entry for nuclear material. Towards that end SAIC and Ludlum PVT radiation portals, based on Pacific Northwest National Labs (PNNL) specifications, have been deployed at most ports of entry. Around ninety-five percent of the containers entering at these ports of entry are currently being inspected. This has enabled collection of terabytes of radiation data on millions of containers and their contents. We propose a new machine learning approach to create a real time decision system to detect illicit nuclear material. Given the magnitude of available data, it should be feasible to implement this approach.

Unified Confidence Bounds for Censored Weibull Data with Covariates

✱Fritz Scholz, University of Washington, Department of Statistics, Box 354322, Seattle, WA 98195-4322, *fscholz@u.washington.edu*

Key Words: quantile, tail probability, monotone, Weibull, bootstrap, maximum likelihood

Data from a 2-parameter Weibull distribution (with covariates and censoring) are usually analyzed on a logarithmic scale because of the location-scale nature of the transformed data. Quantities of interest are the k-vector of regression parameters β, the scale parameter σ, p-quantiles , and tail probabilities for a given threshold y and for any given covariates. Maximum likelihood theory large sample approximations can give confidence bounds for p-quantiles and tail probabilities that are not always monotone in p or y (i.e., are not inverses to each other). This problem is resolved by invoking either the approximate $(k+1)$-variate normal approximation for $((\widehat{\beta\,}-\beta)/\widehat{\sigma},\log(\widehat{\sigma}))$ or its bootstrapped approximating distribution. Accepting these approximations, the remainder proceeds along exact steps.

Incorporating Covariates in Flowgraph Models: Applications to Recurrent Event Data

✱Aparna Huzurbazar, Los Alamos National Laboratory, Statistical Sciences Group CCS-6, MS F 600, Los Alamos, NM 87505, *aparna@lanl.gov*; Brian Williams, Los Alamos National Laboratory

Key Words: Bayesian prediction, system reliability, maintenance, semi-Markov process, data analysis

Modeling recurrent event data is of current interest in statistics and engineering. This article concerns adding covariates to flowgraph analysis with applications to recurrent event data. A flowgraph is a generalized transition graph (GTG) that is used to represent semi-Markov processes. The focus of the flowgraph model is on data analysis for complex stochastic systems. Previous flowgraph modeling did not incorporate continuous-time covariates into the analysis. We consider modeling recurrent event data as arises in the systems reliability setting. Results are presented on simulated and real-data examples. Software for flowgraph modeling is available from the authors.

Goodness-of-Fit Testing from the Pareto and the Logistic Distributions

✱Sneh Gulati, Florida International University, Dept of Statistics,, Florida International University, Univ Park, Miami Beach, FL 33199, *gulati@fiu.edu*; Samuel Shapiro, Florida International University

Key Words: Pareto Distribution, Logistic Distribution, Regression Tests, Greenwood's statistic, Goodness of Fit Test

The Pareto and the Logitic Distributions can serve to model several types of data sets, for example data arising in the insurance industry, reliability studies and life data. In this paper, we present two tests, one to test the hypothesis that the underlying data come from a Pareto Distribution and the other for the Logistic distribution. The tests presented for the Pareto Distributions are based on the regression test of Brain and Shapiro (1983) for the exponential distribution while the one for the Logistic distribution is based on a modified Greenwood Statistic of Chen and Shapiro (2001). Power comparisons of the tests are carried out via simulations.

89 Semi/Nonsupervised Learning and Its Application in Bioscience ●

Biometrics Section, WNAR

Monday, August 4, 8:30 a.m.–10:20 a.m.

Bayesian Nonlinear Classification Methods and Their Applications

✱Veera Baladandayuthapani, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd. Unit 447, HOUSTON, TX 77030, *veera@mdanderson.org*

Key Words: Bayesian, Functional data analysis, microarray, semiparametric methods, MCMC, cancer

This work is directed toward the development of Bayesian nonlinear statistical methods for analyzing high dimensional data from biological experiments. I will discuss two case studies where such methods are particularly useful. The first concerns microarray data. We present a full probabilistic model based approach to nonlinear classification based on adaptive splines. Not only we obtain nonlinear decision boundaries for our classifier but we also model the functional interaction between the genes. The second case study concerns functional data. We present a functional classification model motivated by a real oncology experiment. Flexible semiparametric formulation for modeling the individual curves is proposed via hierarchical regression splines. The inferential framework is essentially Bayesian and uses MCMC methods for estimation.

Constrained Clustering in Linear Array with Applications in Genetics and Microbiology

✱Samiran Ghosh, Indiana University Purdue University, Indianapolis, 402 N. Blackford Street, LD270,, Indianapolis, IN 46202-3216, *samiran@math.iupui.edu*; Jeffrey Townsend, Yale University; Dipak Dey, University of Connecticut

Key Words: Bayesian, Genetics, Hierarchical Clustering, Microbiology, Monte Carlo

Clustering is an ubiquitous data analytic technique. There are several biological domains where the underlying data generating process is constrained. In the present case we are interested in a scenario where data points are distributed in a liner array and they are not exchangeable. In other words each position in the linear array is fixed and could be either “expressed” or “nonexpressed.” We would like to explore whether there exists any kind of clustering among the expressed positions. As data generating process is itself constrained, traditional clustering methods needs suitable adjustment to be successful. In this talk we would like present a Monte Carlo base hierarchical partitioning clustering algorithm, which is specifically suitable for such a constraint domain. We have tested the efficacy of our methods for two biological data sets coming out of genetics and microbiology.

Nonparametric Clustering of Discrete Functional Data

✱Haiyan Wang, Kansas State University, Department of Statistics, 101 Dickens Hall, Manhattan, KS 66506, *hwang@ksu.edu*

Key Words: Clustering, Functional data, Multiple testing, DNA Barcode

Clustering of high dimensional mixtures of discrete functional data will be discussed in this talk. The clusters are defined through the unknown high-dimensional multivariate distributions of all observations from the same subject. Kullback-Leibler information and Mahalanobis distance can fail to provide meaningful measure of distance between distributions in high-dimensional setting. The similarity measure and agglomerative clustering algorithm (PCLUST) in Wang et al. (2008) can effectively differentiate such mixture of high dimensional populations. The algorithm produces invariant results under monotone transformations of data and does not require users to specify the number of clusters. Simulations show that PCLUST significantly outperforms 9 other popular algorithms in clustering accuracy and stability. An application in DNA barcode to identify species will be presented.

Cluster-Analytic Health State Modeling

✱Catherine A. Sugar, University of California, Los Angeles, CHS 51-236C, Department of Biostatistics, School of Public Health, Box 951772, Los Angeles, CA 90095-1772, *csugar@ucla.edu*

Key Words: Cluster Analysis, Health States, Health Policy, Functional Data Analysis

In medical studies, physical and psychological well-being are often measured using health status instruments with dozens of item responses. A small number of continuous summary scores are created and analyzed via standard univariate techniques. In the context of complex diseases this approach may ignore important inter-relationships among dimensions of health. Multivariate health-state models capture the structural richness of such data using clustering combined with Markov chains or functional data analysis to define classes of patients or disease trajectories. This leads to efficient modeling of patient populations, provides natural estimates of long-run treatment effectiveness and facilitates elicitation of utilities for use in public policy analyses. We demonstrate the methods in applications ranging from psychiatry to oncology and discuss some accompanying theoretical issues.

A Mixture Model Approach in Gene-Gene and Gene-Environmental Interactions

✱Lang Li, Indiana University, 410 West 10th St., HITS Suite 3000, Carmel, IN 46032, *lali@iupui.edu*

Key Words: gene-gene interaction, mixture model, pharmacogenetics

In translational research, a genetic association study usually has a two-fold aim: test whether genetic/environmental variables or their combinations are associated with a clinical phenotype; and determine how those combinations are grouped to predict the phenotype (i.e. which combinations have a similarly distributed phenotype, and which ones have differently distributed phenotypes). Although traditional logistic regression can detect gene-gene or gene-environmental interaction effects on binary phenotypes, they cannot decisively determine how genotype combinations are grouped to predict the phenotype. Our proposed mixture model approach is valuable in this context. It concurrently detects main and interaction effects of genetic and environmental variables through a likelihood ratio test, and conducts phenotype cluster analysis based on genetic and environmental variable combinations.

90 Statistical Issues in Medical Device Studies ●

Biopharmaceutical Section, WNAR, Section on Health Policy Statistics, Biometrics Section

Monday, August 4, 8:30 a.m.–10:20 a.m.

Meta-Analysis of Pedicle Screw in Spinal Fusion

✱Harry F. Bushar, U.S. Food and Drug Administration, 1350 Piccard Dr., Rockville, MD 20850, *harry.bushar@fda.hhs.gov*; Hollington T.C. Lu, U.S. Food and Drug Administration

Key Words: meta-analysis, pedicle screw, treatment-covariate interaction, spinal fusion, random effects, meta-regression

Bone screws are used to build and secure spinal constructs and are frequently used with plates, rods, or other devices to affix the components of the superstructure. No prospective, randomized study has been completed which comprehensively evaluates different spinal implants used to correct for spinal trauma. Although, such a study would be ideal, it would be impractical, costly, and require an exceptionally long time to perform. Our meta-analysis study compares the performance of different internal fixation devices for managing and treating thoracolumbar spinal trauma. To achieve this goal, the relevant English-language medical literature was reviewed and the data were subjected to meta-analysis. The method of meta-analysis and the results of different prognostic variables are reported. We propose use of a random-effects meta-regression to study treatment-covariate interaction.

Drug Release Behavior in Drug-Eluting Stents

✱Shanti Gomatam, U.S. Food and Drug Administration, 1350 Piccard Dr HFZ-550, Rockville, MD 20850, *shanti.gomatam@fda.hhs.gov*

Key Words: functional analysis, response surface, medical device

In drug-eluting stents, the drug is incorporated into a polymer matrix, which acts as a diffusion barrier, thus slowing the rate of drug release. The drug-polymer coatings are routinely fabricated by dissolving a mixture of drug and polymer into a solvent. The solvent then evaporates, leaving a system consisting of essentially only drug and polymer. A thermodynamically consistent diffuse interface model has been developed that enables prediction of the structural evolution of controlled release composites during both fabrication and drug release. Calculations based on the theory have been conducted to probe the influence of drug-polymer interactions, drug loading,

and evaporation rate on the effect of release media on delivery kinetics. In this work we utilize these data to assess statistically the impact of variability in the process variables on drug release.

A Permutation Test for a Weighted Measure of Survival Difference with Application to the Stent Thrombosis Examination in Coronary Stenting Trial

✱Jin Wang, Abbott Vascular, 3200 lakeside drive, santa clara, CA 95054, *jin.wang@av.abbott.com*

Key Words: Stent thrombosis, Weighted Kaplan-Meier, Permutation test, Proportional hazards, Drug-eluting stent

For survival analysis between two treatment groups, Kaplan-Meier approach is commonly used with log-rank test being the default test for survival difference. Recent publications suggest many situations that log-rank test may not be appropriate. When a new intervention is to delay a negative effect such as death relative to the control or when the survival curves cross each other, the proportional hazard assumption no longer holds and the log-rank test may suffer substantial power loss. In this paper, two recently published methods are reviewed and a new method is proposed based on modification on one of the methods and comparison among these methods are discussed.

Comparison of Individual Patient-Level and Study-Level Meta-Analyses Using Time-to-Event Analysis in Drug-Eluting Stent Data

✱Hsini Liao, Boston Scientific Corporation, 100 Boston Scientific Way, Marlborough, MA 01752-1234, *terry.liao@bsci.com*; Yun Lu, Boston Scientific Corporation; Hong Wang, Boston Scientific Corporation

Key Words: study-level meta-analysis, individual patient-level data, time to event analysis

Meta-analysis provides a structure of consolidating the outcomes from several studies and deriving statistical inferences of the outcomes. When there is a lack of the individual patient-level data (IPD), the use of a study-level meta-analysis (SMA) can be used to assess the safety and effectiveness of treatment effect. However, it is known that an SMA can lead to biased assessments. Namely, an SMA may or may not agree with an IPD meta-analysis. Here, we focus on the diabetic subgroup in five randomized controlled studies of drug-eluting stents compared to bare metal stents. The hazard ratio estimate of study outcomes from IPD is compared with that from the SMA to assess the treatment effect in diabetic patients. The survival curves of selected clinical outcomes in the intent-to-treat diabetic subgroup are assessed through five years.

Statistical Issues for Combining Clinical Data from Medical Device Multicenter Trials

✱Chang S. Lao, U.S. Food and Drug Administration, 1350 Piccard Dr., Rockville, MD 20850, *chang.lao@fda.hhs.gov*; Harry F. Bushar, U.S. Food and Drug Administration

Key Words: Pooling, Meta-Analysis, Multi-Center Trials

In this paper, we discuss statistical pooling issues and meta-analysis in medical device clinical studies, which include purpose, important conditions, and questions to be addressed in pooling of multi-center studies. Medical device clinical studies may be different from drug trials in terms of study design, overall sample size, and etc. We primarily discuss pooling issues for binary clinical outcomes, in terms of risk difference, risk ratio, and odds ratio, by use of real pre-market approval clinical studies of a tibia bone fracture device. Several statistical issues, such as test for overall treatment effect, homogeneity of treatment effect across centers, combination of multi-center studies, and fixed and random effects models, will be discussed. We conclude that individual-patient clinical data provide more useful information than data from summary studies or published literature.

91 Modern Statistical Machine Learning for Complex and High-Dimensional Data

IMS, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, Section on Statistical Computing, WNAR

Monday, August 4, 8:30 a.m.–10:20 a.m.

Local Quasi-Likelihood Method with a Parametric Guide

✱Yichao Wu, Princeton University, Dept of ORFE, E-418, E-Quad, Princeton University, Princeton, NJ 08540, *yichaowu@princeton.edu*; Jianqing Fan, Princeton University; Yang Feng, Princeton University

Key Words: GLM, local polynomial, quasi-likelihood

Generalized linear models and quasi-likelihood method extend the ordinary regression models to accommodate more general conditional distributions of the response. Despite their popularity, the same deficiency of the parametric ordinary regression applies to parametric generalized linear models in the sense that misspecification of the parametric model can lead to a completely wrong picture of the underlying conditional mean function. In this work, we propose to combine parametric and nonparametric methods. Specifically, two parametrically guided nonparametric estimation schemes are proposed by incorporating some prior shape information into consideration for GLMS and quasilikelihood method. Asymptotic theory and numerical simulations are used to demonstrate their improvement over the original nonparametric method.

Variable Selection in Nonparametric Varying-Coefficient Models for Analysis of Repeated Measurements

✱Lifeng Wang, University of Pennsylvania, 634 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, *lifwang@mail.med.upenn.edu*; Hongzhe Li, University of Pennsylvania; Jianhua Huang, Texas A&M University

Key Words: Regularization, Nonparametric function estimation, Functional shrinkage, Longitudinal data, Time course gene expression data

Nonparametric varying-coefficient models are commonly used for analysis of data measured repeatedly over time. While many procedures have been developed for estimating the varying-coefficients, the problem of variable selection for such models has not been addressed. In this work, we propose a regularized estimation procedure, which simultaneously selects significant variables with time-varying effects and estimates the nonzero smooth coefficient functions. We have established theoretical properties of our procedure, including consistency in variable selection and the oracle property in estimation. The method is illustrated with simulations and two real data examples, one for identifying risk factors in AIDS study and one using microarray time-course gene expression data to identify the transcription factors related to yeast cell cycle process.

High-Dimensional Variable Selection via Nonnegative Garotte

✱Jian Zhang, Purdue University, *jianzhan@purdue.edu*

Key Words: lasso, nonnegative garotte, variable selection

We study the problem of high-dimensional variable selection with some good initial estimator which is element-wise estimation consistent but not variable selection consistent. We show that by using the nonnegative garotte procedure, we can obtain a final estimator which is both efficient in estimation and consistent in variable selection. Our setting is high-dimensional (i.e. we allow the number of variables increase almost as fast as exp(n)). We also study the conditions under which the ridge estimator can be a good initial estimator. Numerical studies are conducted to support the theory.

Feature Selection and Grouping for High-Dimensional Classification Using the OSCAR-SVM

✱Dhruv Sharma, North Carolina State University, 2824-201 Avent Ferry Road, Raleigh, NC 27606, *dbsharma@ncsu.edu*; Howard Bondell, North Carolina State University; Hao H. Zhang, North Carolina State University

Key Words: Classification, Clustering, Collinearity, Grouping, Support Vector Machine, Variable Selection

Variable selection for high-dimensional data analysis offers unique challenges, particularly when there are many redundant noise features and high colinearity. Support Vector Machines have been successfully used to study classification problems. In this paper we propose a new procedure called OSCAR-SVM to perform automatic variable selection for classification problems under such conditions. The proposed method is based on the OSCAR procedure for variable selection in regression problems. The new procedure simultaneously selects variables while grouping them into predictive clusters, further aiding in the investigation of groups of features with similar behavior. The OSCAR-SVM is shown to compare favorably with existing methods in terms of both classification accuracy and dimension reduction, while yielding the additional group structure information.

92 NCHS Survey Programs

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Monday, August 4, 8:30 a.m.–10:20 a.m.

Using the National Health and Nutrition Examination Survey (NHANES) To Evaluate Health Disparities

✱Vicki L. Burt, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *vburt@cdc.gov*; Lester R. Curtin, Centers for Disease Control and Prevention; Cindy Zhang, Harris Corporation

Key Words: national health and nutrition examination survey, nhanes

The NHANES has been conducted since the 1960s and is unique in that physical examination data are obtained. The 1999–2006 NHANES is designed so that estimates of health and nutritional status, risk factors and health conditions can be estimated in age-gender specific domains of the non-Hispanic (NH) black, NH white, Mexican American, and low income populations. In addition to the examination, other unique information collected in NHANES includes dietary intake data, data from household specimens (water and dust), personal exposure monitors and activity monitors, and from tuberculin skin testing. Examples of estimates such as chronic disease prevalence, seroprevalence, exposure levels to environmental chemicals and dietary intake obtained using NHANES data that highlight disparities will be presented Limitations of the NHANES data in making such estimates will be discussed.

Using the National Health Care Surveys To Monitor Disparities in Health Care

✱Nancy Sonnenfeld, National Center for Health Statistics, Division of Health Care Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *nsonnenfeld@cdc.gov*; Jane E. Sisk, National Center for Health Statistics; John E. Orban, Battelle

Key Words: disparities, health care, surveys

The National Health Care Surveys are a family of nationally representative, provider-based surveys that collect data about health-care providers, their patients, and their care. The surveys cover health-care providers across a broad spectrum of ambulatory, hospital, and long-term care settings. The resulting data are using in many ways, including to assess disparities in quality of care and access, and differences in health status, among different U.S. populations. The presentation will include an overview of the surveys and the data that each collects. Examples will be provided of analyses that characterize disparities according to race and ethnicity, expected payment source, geographic location, and age. Initiatives to better assess disparities will be described, including the redesigned National Hospital Discharge Survey.

Vital Statistics: Vital to the Measurement of Health Disparities

✱Charles J. Rothwell, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *CRothwell@cdc.gov*

Key Words: Vital statistics, Life expectancy, Leading causes of death, Infant mortality, Sexual practices

This presentation will provide background on the strengths and weaknesses in the analysis of vital statistics and of data from the National Survey of Family Growth for measuring health outcomes using health disparities as an example. Measures such as infant mortality, life expectancy, low birth weight and prematurity, emerging and rare causes of death, and high risk sexual behaviors will be mentioned. Additional discussion will cover new activities that will have a positive impact on data quality and timeliness.

Using National Health Interview Survey (NHIS) Data and State and Local Area Integrated Telephone Survey (SLAITS) Data To Study Health Disparities

✱Jane F. Gentleman, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *JGentleman@cdc.gov*

Key Words: Health disparities, National Center for Health Statistics, National Health Interview Survey, State and Local Area Integrated Telephone Survey

Examples from published analyses will show the usefulness of data from NHIS and SLAITS for studying health disparities: Married adults were generally healthier than adults in other marital status categories (NHIS); grouping all Asian adults into a single category conceals many differences among Asian subgroups (NHIS); children in step, single-mother, or grandparent-only families had poorer health than children living with two biological parents (SLAITS); the % of adults who were ever tested for HIV increases and then decreases with age (NHIS); the prevalence of diabetes was higher among non-Hispanic black persons and Hispanic persons than among non-Hispanic white persons (NHIS); Hispanic persons were most likely to be without health insurance, followed by non-Hispanic black persons and non-Hispanic white persons (NHIS).

93 Ranked Set Sampling ▲

Section on Nonparametric Statistics, Section on Government Statistics, Section on Survey Research Methods

Monday, August 4, 8:30 a.m.–10:20 a.m.

Applications of Ranked Set Sampling to Genetic Studies

✱Gang Zheng, National Heart, Lung, and Blood Institute, Office of Biostatistics Research, 6701 Rockledge Drive, MSC 7913, Bethesda, MD 20892, *zhengg@nhlbi.nih.gov*; Chen Zehua, National University of Singapore; Kaushik Ghosh, University of Nevada, Las Vegas; Zhaohai Li, The George Washington University

Key Words: ranked set sampling, genetic association, linkage, cost-effective, truncation, simple random sampling

In a typical application, ranked set sampling (RSS) is a cost-effective approach if ranking sampling units is easier and cheaper than measuring them. In this case, RSS is useful compared to classical random sampling. In genetic studies, genotypes and phenotypes (traits) of individuals have to be measured and obtained in genetic association and linkage studies. In many applications, genotyping cost is much more expensive than measuring phenotypes. Therefore, RSS can be applied. In this talk, we present two examples, where RSS can be easily applied, compared to using random sampling and the truncation approach. The first example is to map linkage disequilibrium and the second example is to test linkage using affected sib-pairs. In both examples, the idea of extreme rank selections will be presented. Simulation results will also be presented to compare different approaches.

Some Further Generalizations of Ranked Set Sampling

✱Kaushik Ghosh, University of Nevada, Las Vegas, Department of Mathematical Sciences, 4505 Maryland Parkway, Box 454020, Las Vegas, NV 89154-4020, *kaushik.ghosh@unlv.edu*; Ram C. Tiwari, National Cancer Institute

Key Words: Empirical process, P-P plots, Q-Q plots, Extreme ranked set sample

We generalize the idea of ranked set sampling by allowing it to incorporate dependent order statistics of various sizes. The data so obtained is used to estimate the distribution function and compare two distributions. Asymptotic results are established and are compared with simulation runs. An application using weather data is also provided.

Nonparametric Maximum Likelihood Estimator of Bohn-Wolfe Model

✱Omer Ozturk, The Ohio State University, 404 Cockins Hall, 1958 Neil aven, Columbus, OH 43210, *omer@stat.osu.edu*

Key Words: Imperfect ranking, ranked set sampling, Judgment ranking, calibration

Many inferential procedures based on a ranked set sample data do not allow the possibility of ranking error. On the other hand, in a typical application, ranking error may not be eliminated and statistical inference may be impacted by the severity of this error. In this talk, we introduce a nonparametric maximum likelihood estimator for Bohn and Wolfe (1994) judgment ranking model. It is shown that the MLE of the cumulative distribution function (cdf) of the underlying population exist for any legitimate Bohn-Wolfe model. We also show that the MLE of the judgment ranking probabilities exist for any legitimate CDF estimator. We estimate the Bohn-Wolfe model by iterating between these two estimators until we have a convergence. The properties of the estimators are investigated in a simulation study. The estimated Bohn-Wolfe model is used to calibrate the impact of imperfect ranking.

Confidence Intervals Estimation of the Location Parameter of the Logistic Distribution Using Ranked Set Sampling

✱Hassen A. Muttlak, KFUPM, KFUPM, P.O. Box 1676, Dhahran, 31261 Saudi Arabia, *hamstat@kfupm.edu.sa*

Key Words: efficiency, length, maximum likelihood estimator, ranked set sampling, simple random sampling, standard error

The maximum likelihood estimator (MLE) and several different pivots will be used to construct different confidence intervals for the logistic location parameter. These confidence intervals will be constructed by using simple random sampling (SRS) and ranked set sampling (RSS). Difference confidence intervals will be compared via their expected lengths and the standard deviation of their lengths using simulation. We expect that the confidence intervals base on RSS to have shorter expected lengths and smaller standard deviation from there competitors using the SRS.

Sequential Unbalanced Ranked Set Sampling

✱Nader M. Gemayel, The Ohio State University, 404 Cockins Hall, 1958 Neil Ave, Columbus, OH 43210, *nader@stat.osu.edu*; Douglas Wolfe, The Ohio State University; Elizabeth Stasny, The Ohio State University

Key Words: Sequential methods, Ranked set sampling, population proportion

One is often interested in estimating the population proportion of a rare event, where conventional sampling techniques (such as SRS and balanced RSS) are wasteful and inefficient. Moreover, one is also often interested in the units of a population that exceed a certain elevated threshold; for instance, the proportion of fraudulent accounts in a population of accounting records, and the distributional properties of the largest amounts of dollar fraud among the fraudulent accounts. We will explore how sequentially-guided RSS techniques can lead us to target sparsely distributed occurrences of the rare event more accurately.

94 Value-Added Models for Student Achievements ●

Social Statistics Section, Section on Survey Research Methods

Monday, August 4, 8:30 a.m.–10:20 a.m.

Value-Added and Test-Score Ceiling Effects

✱Cory R. Koedel, University of Missouri, 909 University Avenue, 327 Professional Building, Columbia, MO 65211, *koedelc@missouri.edu*; Julian R. Betts, University of California, San Diego

Key Words: value added, teacher quality, test score ceiling, teacher effects

This paper uses administrative data to evaluate the importance of test-score ceiling effects in the estimation of value-added. There is no evidence of an actual test-score ceiling in the raw data, which is particularly appealing for this project. Starting with the no-ceiling baseline, we consider a variety of artificially imposed test-score ceilings in a variety of estimation conditions. We determine the extent to which test-score ceilings affect teacher and school rankings based on value-added and whether they help or harm particular types of teachers and schools. We also evaluate the effects of test-score ceilings on the precision of value-added estimates and the overall measurable variance of teacher and school quality.

The Sensitivity of Teacher Effect Estimates to Decision Rules for Establishing Student-Teacher Links

✱Matthew Springer, Peabody College of Vanderbilt University, *alicen.i.hatter@vanderbilt.edu*

One dimension untested in value-added (VA) research is the sensitivity of teacher effect estimates to rules for defining student-teacher links. Student-teacher links are complicated by an array of factors, including students transferring among schools mid-year; transferring between classes mid-year; taking multiple same-subject courses; or being habitually absent. Because performance measures may have high stakes in merit pay plans rule changes that result in small adjustments in student-teacher links may have a significant impact for some teachers. Our study fills this gap by: reviewing how VA studies, to date, have defined student-teacher links; creating a taxonomy that captures rules used when defining student-teacher links; and conducting a simulation study that examines the sensitivity of teacher VA estimates to decision rules made when defining student-teacher links.

Does a Teacher's Value Added Require Data To Be Missing at Random?

✱Daniel McCaffrey, RAND Corporation, 4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665, *daniel_mccaffrey@rand.org*; J. R. Lockwood, RAND Corporation

Key Words: Bayesian models, Mixed models, Teacher effects, Variable Persistence Model, Student Achievement

Incomplete data are a common concern for analyses that estimate teacher contributions to learning from longitudinal student achievement data. Students with missing scores are at greater risk for low performance and data might not be missing at random (MAR). We develop models for estimating teacher effects that allow for missing not at random data. We consider a selection model where the number of observed test scores depends on a student's latent general level of achievement and a pattern mixture model that allows the means and covariances of test scores to depend on the student's pattern of observed scores. We fit these models to five years of longitudinal test score data from a large urban school district. These models yield very similar teacher effects as models that assume data are MAR. We discuss likely explanations for the robustness of estimates to assumptions about missing data.

Estimating Teacher Effects from Longitudinal Data Without Assuming Vertical Scaling

✱Louis T. Mariano, RAND Corporation, 1200 S. Hayes St, Arlington, VA 22202, *Lou_Mariano@rand.org*; Daniel McCaffrey, RAND Corporation; J. R. Lockwood, RAND Corporation

Key Words: value added methods, vertical scaling, Bayesian hierarchical models, MCMC, Wishart prior, multiple-membership models

There is increasing interest in using longitudinal measures of student achievement to estimate individual teacher effects. Current multivariate models assume scores come from a single vertical scale and that each teacher has a single effect on student outcomes that persists undiminished to all future test administrations (complete persistence) or can diminish with time but remains perfectly correlated (variable persistence). However, vertically linked tests might not be unidimensional, and not all state assessments use a vertical scale. We develop the "generalized persistence" model, a Bayesian multivariate model for estimating teacher effects that accommodates longitudinal data that are not vertically scaled by allowing a teacher's effects on her student's current and future outcomes to have less than perfect correlation. We illustrate the model using mathematics assessment data.

Estimating Dynamic Panel Data Models with Measurement Errors

✱Shawn Ni, University of Missouri, Department of Economics, University of Missouri, Columbia, MO 65211, *NiX@missouri.edu*; Michael Podgursky, University of Missouri

Key Words: measurement errors, school effect

A key difficulty in drawing inference of school effect from student test score gains is the fact that test scores are noisy measurements of students' academic achievements. In this paper we examine competing inferential methods of dynamic panel data models using noisy data. In particular, we consider a score-level model where the score of the current period depends on the score of the previous period. We compare Monte Carlo simulation results of four estimators of these models and obtain these estimates using the student test score data of Missouri.

95 Student Paper Competition: Bayesian Methods

Section on Bayesian Statistical Science

Monday, August 4, 8:30 a.m.–10:20 a.m.

A Bayesian Hierarchical Model with Curve Selection for Functional Data Classification

✱Hongxiao Zhu, Rice University, 6100 Main St. MS-138, Houston, TX 77005, *hxzhu@stat.rice.edu*; Marina Vannucci, Rice University; Dennis D. Cox, Rice University

Key Words: Bayesian hierarchical model, functional data classification, curve selection, Evolutionary Monte Carlo, fluorescence spectroscopy

In functional data classification, functional observations are often contaminated by various systematic effects, such as random batch effects caused by device artifacts, or fixed effects caused by sample-related factors. Another issue of concern is the selection of functions when the predictors consist of multiple functions with redundancy, which we call "curve selection." In this paper, we use a Bayesian hierarchical model to take into account random batch effects, and propose a block-wise variable selection method to choose effective curves among multiple functional predictors. Fixed effects or predictors in non-functional form are also included in the model. A hybrid Metropolis-Hastings/Gibbs sampler is used for posterior sampling, and an Evolutionary Monte Carlo algorithm is applied to improve the mixing property of the curve selection parameter.

Asymptotic Comparisons of Predictive Densities for Dependent Observations

✱Xuanyao He, The University of North Carolina at Chapel Hill, Dept of STOR, Chapel Hill, NC 27599-3260, *xoyo@unc.edu*; Richard L. Smith, The University of North Carolina at Chapel Hill; Zhengyuan Zhu, The University of North Carolina at Chapel Hill

Key Words: Mixed effect models, Kullback-Leibler divergence, Jeffreys prior, Predictive density, Prediction fit

This paper studies Bayesian predictive densities based on different priors and frequentist plug-in type predictive densities when the predicted variables are dependent on the observations. Average Kullback-Leibler divergence to the true predictive density is used to measure the performance of different inference procedures. The notion of second-order KL dominance is introduced, and an explicit condition for a prior to be second-order KL dominant is given using an asymptotic expansion. As an example, we show theoretically that for mixed effects models, the Bayesian predictive density with prior from a particular improper prior family dominates the performance of REML plug-in density, while the Jeffrey's prior is not always superior to the REML approach. Simulation studies are included which show good agreement with the asymptotic results for moderate sample size.

On Bayesian Inference for Generalized Multivariate Gamma Distributions

✱Sourish Das, University of Connecticut, 215 Glenbrook Road, Department of Statistics, Storrs, CT 06269, *sourish.das@gmail.com*; Dipak Dey, University of Connecticut

Key Words: Autoregressive structure, Bayesian Inference, Dispersion Matrix, MAP estimate, Model Selection

In this paper, we define a generalized multivariate gamma distribution with infinite divisibility property. Then, we show the different patterned dispersion matrix (e.g., compound symmetry structure, autoregressive structure), which satisfies certain infinite divisibility condition. We develop the important properties of the multivariate gamma distribution. Then, we consider Bayesian decision theoretic approach to develop the Bayesian inference for multivariate gamma distribution.

Efficiently Estimating Personal Network Size

✱Tyler H. McCormick, Columbia University, Room 1005, MC 4690, Department of Statistics, 1255 Amsterdam Ave, New York, NY 10027, *tyler@stat.columbia.edu*; Matthew Salganik, Princeton University; Tian Zheng, Columbia University

Key Words: social networks, estimating degree, latent non-random mixing model

In this paper we develop a method to estimate both individual social network size (i.e., degree) and the distribution of network sizes in a population by asking respondents how many people they know in specific subpopulations (e.g., people named Kevin). Building on the scale-up method of Killworth et al.(1998) and other previous attempts, we first propose a latent non-random mixing model which resolves three known problems with previous approaches. As a byproduct, our method also provides estimates of the rate of social mixing between population groups. We demonstrate the model using a sample of 1,370 adults originally collected by McCarty et al. (2001). We conclude by offering practical guidelines for the design of future surveys in this area. Most importantly, we show that if the specific subpopulations are chosen wisely, complex statistical procedures are no longer required.

Another Look at Macroeconomic Forecasting: Aggregation vs. Disaggregation

✱Kun Ho Kim, The University of Chicago, Dept of Economics, 4060 Springer Way, Apt 1117, East Lansing, MI 48823, *kunhokim@uchicago.edu*

Key Words: Macroeconomic forecasting, Disaggregation, Reduced-form equation model, Hierarchical model, Shrinkage effect, Bayesian MCMC

In this paper, the "Revised Expectations Model" (REM), a variant of the Marshallian Macroeconomic Model (MMM) is developed to incorporate agents' price expectation formation effects. With this modification incorporated in an aggregate one sector and in a disaggregated multi-sector version of the MMM, the aggregate and disaggregate models are estimated and used in forecasting experiments. These experiments show that use of the disaggregated version of the model incorporating price expectation effects along with modern Bayesian MCMC estimation and prediction techniques produce more precise predictions than those yielded by an aggregate version and by several benchmark forecasting models. It appears that the REM will be useful not only in forecasting economic variables but also in analyzing the effects of policy changes on those variables.

96 Best Practices in Statistics Training: Lessons Learned from VIGRE Programs

Section on Statistical Education
Monday, August 4, 8:30 a.m.–10:20 a.m.

Best Practices in Statistics Training: Lessons Learned from VIGRE Programs

✱Tilmann Gneiting, University of Washington, Department of Statistics, Seattle, WA Washington 98195-432, *tilmann@stat.washington.edu*; ✱Susan Holmes, Stanford University, Statistics- Sequoia Hall, 390 Serra Hall, Stanford, CA 94305, *susan@stat.stanford.edu*; ✱Robert Kass, Carnegie Mellon University, *kass@stat.cmu.edu*; ✱Sastry Pantula, North Carolina State University, *pantula@stat.ncsu.edu*; ✱John Rice, University of California, Berkeley, *rice@stat.berkeley.edu*

Key Words: education, curriculum revision, broader impact, VIGRE, professional development, undergraduate research

The NSF VIGRE program supports activities involving an entire department, with the goal of enhancing the education of undergraduates, graduates, and postdocs. The award requires: broad faculty involvement; interaction between undergrads, grads, postdocs, and faculty; curriculum review of undergrad and grad programs; procedures for mentoring and advising; research opportunities for undergrads; professional development for grad students; and significant efforts to increase the numbers for under-represented groups in statistics. Over the past eight years, VIGRE-supported departments have seen many positive changes in their programs, and they have abandoned unsuccessful attempts to improve their programs. Faculty from VIGRE-supported departments will present the successes, failures, and significant changes to their departments, and indicate what is transferable to other programs.

97 Statistical Disclosure Control

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Monday, August 4, 8:30 a.m.–10:20 a.m.

Translating Disclosure Risk of Geographic Units to Survey Respondents: A Hierarchical Assessment of Contextual Data

✱Kristine M. Witkowski, The University of Michigan, P.O. Box 1248,Institute for Social Research, Ann Arbor, MI 48106-1248, *kwitkow@umich.edu*

Key Words: confidentiality, contextual, disclosure, geographic, reidentification, spatial

Contextualized microdata is one way to safely release geographic data without identifying the location of survey respondents. This study informs the design of such datafiles with its needle-in-haystack approach to disclosure and its discussion of associated methodological concerns. Drawing a sample of counties, tracts, and blockgroups, I illustrate how the reidentification of individuals is shaped by aggregating geographies into look-alike sets. I detail the complexity of reidentification patterns by assessing the likelihood that young adult white and black males would be pinpointed within reconstituted haystacks given: (1) the size of the total population of aggregated contexts; (2) the amount of error in population counts; and (3) differential search costs stemming from spatially dispersed contexts.

A Multiple Imputation Approach to Disclosure Limitation for High-Age Individuals in Longitudinal Studies

✱Di An, Merck & Co., Inc., *di_an@merck.com*; Roderick J. Little, The University of Michigan

Key Words: confidentiality, disclosure protection, longitudinal data, multiple imputation, survival analysis

Disclosure limitation is an important consideration in the release of public use data sets. It is particularly challenging for longitudinal data sets, since information about an individual accumulates over time. We consider problems created by high ages in cohort studies. Because of the risk of disclosure, ages of very old respondents can often not be released, as stipulated by the Health Insurance Portability and Accountability Act (HIPAA). Top-coding of individuals beyond a certain age is a standard way of dealing with this issue, but it has severe limitations in longitudinal studies. We propose and evaluate an alternative to top-coding for this situation based on multiple imputation (MI). This MI method is applied to a survival analysis of simulated data and data from the Charleston Heart Study, and is shown to work well in preserving the relationship between hazard and covariates.

The Effect of Measurement Error and Under-Coverage Error on Disclosure Risk Assessments

✱Mandi Yu, The University of Michigan, 426 Thompson ST, Room4050, Ann Arbor, MI 48106, *mandiyu@isr.umich.edu*; Trivellore Raghunathan, The University of Michigan

Key Words: Disclosure Risk, Measurement Error, Re-identification, Record Linkage

Conventional wisdom suggests that the availability of commercial databases with identifying information and key demographic variable coupled with powerful record linkage techniques would increase the risk of disclosure of survey respondent. The objective of this paper is to address these concerns by evaluating such risk of disclosure using national surveys and also assess the accuracy of information in the commercial database. The results from these experiments suggest that the disclosure threat from linking entities in the commercial list with those in public data may be exaggerated. Any assessment of risk of disclosure and procedures used to treat the data should take into account the data accuracy used by potential intruder.

Combining Statistical Disclosure Limitation Methods To Create Synthetic Data

✱Jennifer C. Huckett, Iowa State University, 2625 N Loop #2140, Ames, IA 50011, *jhuckett@gmail.com*; Michael D. Larsen, Iowa State University

Key Words: statistical disclosure limitation, hot deck imputation, rank swapping, quantile regression, disclosure risk, data utility

Government agencies face demands to release accurate, timely data while upholding respondents' privacy and confidentiality. We consider combining methods to produce releasable, usable data. Conditional quantile regression models can capture complex relationships characteristic of demographic and economic data. Values for several confidential variables are simulated using predicted values at randomly selected quantiles. Additional values are imputed from original records using hot deck imputation and rank swapping. Combining variable values produces a microdata set for release with low disclosure risk and high data utility. Developments in this work include simulation methods, methods of measuring disclosure risk and measuring for inferential purposes. Applications to U.S. Census Bureau Veterans data and Iowa Department of Revenue Income Tax Returns are discussed.

Achieving Analytic Potential While Protecting Data Confidentiality

Fritz Scheuren, The University of Chicago; ✱James Sears, U.S. Social Security Administration, Washington, DC , *jim.sears@ssa.gov*

In this paper, we follow the lead of Don Rubin, Jerry Reiter, Arthur Kennichell, and Julia Lane, who have recommended a synthetic data set be created—one that produces the same key analytic results as the original data set. By this, we mean in a way to be defined in detail that the original and synthetic data sets are sufficiently close statistically for the synthetic data set to be an "analytically sufficient" substitute. We do not offer a general solution to this proposition. Instead, we discuss the degree of "success" of one such effort, involving a linkage of Survey of Income and Participation (SIPP) data with benefit and earnings data from the Social Security Administration and the Internal Revenue Service, respectively. We look at the analytic strengths and weaknesses of the resulting synthetic data set.

A Unified Theory of Randomized Response Surveys of Binary Characteristics

✱Tapan K. Nayak, The George Washington University, Department of Statistics, George Washington University, Washington, DC 20052, *tapan@gwu.edu*

Key Words: Respondents' privacy, Efficiency, Unbiasedness, Admissibility

Many randomized response (RR) procedures for surveying sensitive binary variables are available in the literature. We shall present a common framework for all RR surveys of binary variables. The unified approach is focused on the substantive issues relating to respondents' privacy and statistical efficiency and is helpful for fair comparison of different procedures. We shall describe an approach for comparing RR procedures, taking both respondents' protection and statistical efficiency into account. For any given RR procedure with three or more response categories, we shall show that one can design an RR procedure with a binary response variable that is as protective and at least as informative. This result suggests that one should use only binary response variables when designing an RR survey of a dichotomous population. Finally, we shall characterize all admissible RR designs.

Randomized Response Model in a Matched Pair Study

✱Chien-Hua Wu, Chung-Yuan Christian University, 32023 Taiwan, *cwu@cycu.edu.tw*

Key Words: Matched Pair, Randomized Response Model

The development of randomized response models for personal interview surveys has attracted much attention since the pioneering work of Warner (1965). Several randomized response models have been developed by researchers for collecting data on both qualitative and the quantitative variables, but none of these models discuss matched pair data. In this paper, we develop a new randomized response model and study its application to an important political question.

98 Nonresponse Bias Analysis ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Monday, August 4, 8:30 a.m.–10:20 a.m.

Nonresponse Bias in the American Time Use Survey

✱Phawn Letourneau, U.S. Census Bureau, *phawn.m.letourneau@census.gov*; Andrew Zbikowski, U.S. Census Bureau

Key Words: Survey, Nonresponse, Logistic Regression

The American Time Use Survey (ATUS) provides estimates of how people divide their time each day in the US. Households that have completed their eighth and final monthly interview in the Current Population Survey (CPS) are eligible for selection in the ATUS. Since the initial data collection in 2003, the ATUS response rate has been below 60%. The low response rate has caused concern over the potential for bias in the ATUS estimates. In 2005, the Bureau of Labor Statistics conducted research on nonresponse in the 2004 ATUS. We follow with an analysis of nonresponse on the 2006 ATUS. We link the ATUS weighted output files to the call history files for nonresponse unit information, and link to the CPS files for available frame variables. We use weighted or survey logistic regression to model response propensities in the ATUS for individuals or households with different characteristics.

Using a Large-Scale Field Study To Estimate Nonresponse and Noncoverage Biases in an RDD Survey

✱Wendy VanDeKerckhove, Westat, 1650 Research Blvd., Rockville, MD 20850, *WendyVanDeKerckhove@Westat.com*; Jill Montaquila, Westat; J. Michael Brick, Westat

Key Words: CATI, Response rate

Concerns about declining response and landline telephone coverage rates in Random Digit Dial (RDD) surveys led to the development and administration of an independent bias study that was conducted in conjunction with the 2007 National Household Education Surveys Program (NHES:2007). For the bias study, a clustered sample of addresses was selected, telephone numbers were matched to the addresses to the extent possible, and the matched sample used the regular RDD protocol. Addresses that could not be matched or completed by telephone were sent into the field for in-person attempts. In this paper, we describe the methods used to estimate nonresponse and noncoverage bias, present estimates of the two types of bias, and interpret the findings in the broader context of other nonresponse and noncoverage bias studies.

A Study of SES Bias and Nonresponse in a Large-Scale Household Survey

✱Orlaith Burke, University College Dublin, School of Mathematical Sciences, James Joyce Library Building, UCD, Belfield, Dublin, International Dublin 4 Ireland, *orlaith.burke@ucd.ie*; Patrick E. Murphy, University College Dublin; Richard Sinnott, University College Dublin

Key Words: SES bias, Non-response, Household surveys

The widely reported phenomenon of socio-economic status (SES) bias in response rates to surveys remains one of the major problems which researchers face when dealing with survey data. The purpose of this paper is to present the results of a study, conducted in Ireland, which provided an unprecedented opportunity to examine the characteristics of both the nonrespondents and respondents to a survey and to compare characteristics of both with the target population. Anonymized records on 41,280 participants in the Irish Quarterly National Household Survey (a nationwide survey of households conducted by the Irish Central Statistics Office) provided data for this study into SES bias. Our findings appear to contradict the perceived position on SES bias in surveys.

Identifying Nonresponse Bias Using a Survey Subset in Follow-Up

✱Harper Gordek, RTI International, Cox 118, 3040 Cornwallis Rd, Durham, NC 27713, *gordek@rti.org*; Patrick Chen, RTI International; Judith Lynch, RTI International; Thomas Morgan, RTI International

Key Words: nonresponse bias, follow-up survey, telephone survey, mail survey

Nonresponse bias occurs when the answers of respondents are different from the presumed answers of nonrespondents. Its analysis is critical for surveys with low response rates. Usually, the analysis is limited to adjusting for information gathered in the early stages of data collection. To study nonresponse for a recent inpatient satisfaction survey, RTI conducted a telephone follow-up using a supplement of survey questions. We then implemented three approaches: comparing survey responses for sample frame attributes with significantly different response rates, comparing survey responses to the second questionnaire mailing with those for the first mailing and to the overall data, and comparing survey outcome variables between mail survey respondents and mail survey nonrespondents. The results from these three approaches provided strong evidence for the presence of nonresponse bias.

Sensitivity Analysis of Nonresponse Bias in the Current Population Survey

✱John Dixon, Bureau of Labor Statistics, 10308 Meredith Ave, Kensington, MD 20895, *dixon_j@bls.gov*

Key Words: nonresponse bias, sensitivity analysis, survey

Survey nonresponse bias is difficult to assess since the difference between responders and nonresponders are rarely measurable. Sensitivity analysis is increasingly used to explore the impact of nonresponse. Date from the Current Population Survey are used to study different types of sensitivity analysis in reporting the potential for bias.

Proxy Pattern-Mixture Analysis for Survey Nonresponse

✱Rebecca Andridge, The University of Michigan, Department of Biostatistics, School of Public Health, Ann Arbor, MI 48109, *fedarko@umich.edu*; Roderick J. Little, The University of Michigan

Key Words: nonignorable nonresponse, missing data, survey data, Bayesian methods

We consider assessment of nonresponse bias for the mean of a survey variable Y subject to nonresponse. We assume that there are a set of covariates observed for nonrespondents and respondents. To reduce dimensionality and for simplicity we reduce the covariates to a proxy variable X that has the highest correlation with Y, estimated from a regression analysis of respondent data. We consider adjusted estimators of the mean of Y that are maximum likelihood for a pattern-mixture model with different means of Y and X for respondents and nonrespondents, assuming missingness is an arbitrary function of a known linear combination of X and Y. We propose a sensitivity analysis, sketch Bayesian versions of this approach, and propose a taxonomy for the evidence concerning bias based on the strength of the proxy and the deviation of the mean of X for respondents from the overall mean.

Analyses of Nonresponse Bias Resulting from Failure To Obtain Consent To Contact Vaccination Providers and Provider Nonresponse in the National Immunization Survey

✱Benjamin Skalland, National Opinion Research Center, *skalland-benjamin@norc.org*; Robert Montgomery, National Opinion Research Center; Philip J. Smith, Centers for Disease Control and Prevention

Key Words: Nonresponse, Bias, NIS

The National Immunization Survey is a national, list-assisted RDD survey conducted by NORC for CDC and monitors vaccination rates of children between 19 and 35 months of age. It consists of two data collection phases: a telephone survey to identify households with age-eligible children followed by a mail survey of providers to collect provider-reported vaccination histories from those children. At the end of the telephone survey consent is requested to contact vaccination providers. We explore potential bias in estimated vaccination coverage rates attributable to (i) the failure to obtain consent at the end of the telephone survey and (ii) provider nonresponse in the provider mail survey.

99 Environmental Statistics

Section on Statistics and the Environment, Section on Risk Analysis

Monday, August 4, 8:30 a.m.–10:20 a.m.

Wind Forecasting Models and Loss Function

✱Amanda S. Hering, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143, *mhering@stat.tamu.edu*; Marc G. Genton, University of Geneva

Key Words: Directional data, Loss functions, Skew-t distribution

Making accurate predictions of wind speed reduces the risk that electrical grids face after accepting wind energy as a source. Two models for making short-term forecasts are presented that handle the wind vector as either Polar or Cartesian coordinates. Modeling the wind vector with speed and direction is more intuitive but also more difficult since direction is a circular variable. Transforming these into the Cartesian components is commonly done, and a regression model with a skew-t error distribution can provide a flexible fit to the data. The quality of the predictions from these models can be more realistically assessed with a loss function that depends upon the "power curve" that describes the nonlinear relationship between speed and power. The proposed loss function yields more insight into the true worth of each model's predictions.

Probabilistic Wind Forecasting Using Bayesian Model Averaging

✱J. McLean Sloughter, University of Washington, 2600 Eastlake Ave E #204, Statistics, Seattle, WA 98102, *mclean@stat.washington.edu*

Key Words: ensemble, forecasting, BMA, wind

Bayesian model averaging (BMA) is a statistical way of post-processing forecast ensembles to create predictive PDFs for weather quantities. It represents the predictive PDF as a weighted average of PDFs based on the individual forecasts. It was developed initially for quantities whose PDFs can be approximated by normal distributions, such as temperature and sea level pressure, and has been extended to skewed distributions for modeling precipitation. In this study BMA is extended to another case of skewed distributions for modeling wind speed. The method was applied to daily 48-h forecasts of maximum wind speed in the North American Pacific Northwest in 2003-04 using the University of Washington mesoscale ensemble, and is shown to provide calibrated and sharp probabilistic forecasts. Comparisons are made between a number of potential formulations for the BMA model.

Test of Homogeneity in Some Cyclic Predator-Prey Systems

✱Severien Nkurunziza, University of Windsor, 401 Sunset Avenue, Windsor, ON N9B 3P4 Canada, *severien@uwindsor.ca*

Key Words: likelihood ratio test, Interaction parameter, Ornestein-Uhlenbeck processes, Lotka-Volterra system, ODE system, predator-prey

In this talk, we consider a stochastic model where k deterministic Lotka-Volterra systems of ordinary differential equations (ODEs) are perturbed with k pairs of random errors. Further, we are interested in testing problem concerning the homogeneity between k pairs of the interaction parameters of the ODEs. We assume that the k pairs of random errors are independent and that, each pair follows correlated Ornstein-Uhlenbeck processes. This stochastic model extends that suggested in Froda and Colavita (2005) as well as in Froda and Nkurunziza (2007), where k=1. Under this statistical model, we propose a likelihood ratio test and study the asymptotic properties of this test.

Introducing the S-Value: An Exploratory Tool for Spatial Dependence on a Lattice

✱Petrutza Caragea, Iowa State University, Snedecor Hall, Ames, IA 50014, *pcaragea@iastate.edu*; Mark Kaiser, Iowa State University

Key Words: CAR models, lattice data, exploratory tool

The application of Markov random fields to problems involving spatial data on lattice systems (as often desirable in the environmental and ecological sciences, agriculture, and other areas of biology) requires decisions regarding a number of important aspects of model structure. Existing exploratory techniques appropriate for spatial data do not provide direct guidance to an investigator about these decisions. We introduce a diagnostic quantity useful in situations for which one is contemplating the application of a Markov random field model based on conditional one parameter exponential family distributions. This exploratory diagnostic is shown to be a meaningful statistic that can inform decisions involved in modeling spatial structure with statistical dependence terms. We illustrate its use in guiding modeling decisions and demonstrate that these properties have use in applications.

Adapting a Local Indicator of Spatial Association To Identify Hot Spots in Traffic Safety

✱Elke A. Moons, Transportation Research Institute - Hasselt University, Science Park 1/15, Diepenbeek, International 3590 Belgium, *elke.moons@uhasselt.be*; Tom Brijs, Transportation Research Institute - Hasselt University; Geert Wets, Transportation Research Institute - Hasselt University

Key Words: spatial statistics, local indicator of spatial association, Moran's I, hot spot analysis, traffic safety

In the last years, traffic safety has become a 'hot' topic in the media, for policy makers, for academics and for the broad audience. In general, traffic safety analysis can be split up in four phases: identification of hot spots, ordering, profiling of the hazardous locations and finally selecting the locations that need to be handled. This paper will focus on phase 1. Several spatial measures have been developed to identify hot spots, however, not all measures are equally well suited to the field of traffic safety (e.g. a high-accident location is always of interest for hot spot analysis, but a low-accident location may be of interest if its neighboring locations show an opposite behavior regarding the number of accidents). In this paper we will present an adaptation of a local measure of spatial association, more specific on Moran's I and apply it to identify hot spots on highways.

Mixture Methods for the Analysis of Flood Frequency Distributions

✱John Grego, University of South Carolina, Statistics Departmenbt, University of South Carolina, Columbia, SC 29208, *grego@stat.sc.edu*; Philip Yates, California State Polytechnic University, Pomona

Key Words: EM algorithm, Acceleration methods, Gumbel distribution, Delta method, Edgeworth approximation

Hydrologists have long used unmixed distributions to estimate flood frequency distributions, and have recently proposed improvements to their methods, including use of extreme value distributions. This presentation looks at alternative approaches to this type of analysis to account for multiple mechanisms for flood events. We study finite mixture models for 3-parameter log gamma (log Pearson Type III), normal, lognormal and extreme value distributions. These techniques will then be used on flood frequency data from watersheds in South Carolina to estimate, e.g., the 1% chance flood for these estuaries. We will review standard error estimates for these methods and propose new approaches.

100 Statistics as Liberal Arts, Communication, Critical Thinking, and Outreach ▲

Section on Statistical Education
Monday, August 4, 8:30 a.m.–10:20 a.m.

Teaching Statistics as Part of a Learning Community

✱Yvonne M. Zubovic, Indiana University Purdue University, Fort Wayne, 2101 Coliseum Blvd., East, Fort Wayne, IN 46805, *zubovic@ipfw.edu*

Key Words: education, introductory statstics, writing

Curricular learning communities are classes that are linked thematically during an academic term and enroll a common cohort of students. As part of the First Year Experience initiative, a general education statistics course was linked with a writing course for a cohort of first semester freshmen. We will discuss the learning objectives, for the learning community and the statistics course itself, as well as strategies used to achieve these objectives. We will also present some issues relevant to teaching in a learning community: identifying an interdisciplinary theme, implementing community building activities, coordinating the schedules of topics, integrating joint activities and assignments, designing joint assessments, and organizing co-curricular activities. This talk will highlight lessons learned from teaching an introductory statistics course as part of a learning community.

SiSSYS: Reflections from a Capstone Course Using Sports Data

✱Michael Schuckers, St. Lawrence University, Dept of Math, CS and Stats, Canton, NY 13617, *schuckers@stlawu.edu*; Robin Lock, St. Lawrence University; Travis Atkinson, St. Lawrence University

Key Words: sports, capstone, education

Topics in sports are often used as a conduit for student interest in statistics and statistical methods. In this talk we describe a capstone course based on applications of statistics methods to sports. The Statistics in Sports Senior Year Seminar (SiSSYS) was run at St. Lawrence University during the 2007–08 academic year and included eight senior students. We will include a discussion of the impetus for the seminar as well as a synopsis of the structure, examples of topics considered and an assessment of the outcomes. Additionally, we will make recommendations and offer advice to others considering sponsoring such a course. Two posters at this JSM will display summaries from the final projects that students in this course completed.

Social Justice, Gender Equity, and Service Learning in Statistics Education: Lessons Learned from the DOE-Funded Project ACE

✱Lawrence M. Lesser, The University of Texas at El Paso, Dept. of Mathematical Sciences, 500 W. University Avenue, El Paso, TX 79968-0514, *Lesser@utep.edu*

Key Words: social justice, gender, education, equity, service learning, introductory

Social justice, equity, and service learning are powerful vehicles for supporting statistics education (& the JSM 2008 theme). With support from Project ACE (funded by the U.S. Department of Education; PI: J. Tinajero), I applied the groundwork of my 2007 JSE paper and CAUSE webinar by redesigning an introductory statistics course I taught in fall 2007 to 52 pre-service teachers at a mid-sized research university on the US-México border. We'll discuss the significant quantitative and qualitative evidence of impact on students' knowledge and engagement. We'll also share examples of connecting ACE themes with standard course topics, as well as practical implementation issues. Research suggests that all instructors (regardless of one's personal or political beliefs) need to be aware of the implicit role equity plays in how students experience certain ideas (e.g., bias, random selection).

Interesting Examples of the Use and Misuse of Statistics from the Law

✱Harry J. Norton, Carolinas Medical Center, ROB #410, PO Box 32861, Charlotte, NC 28232-2861, *jnorton@carolinas.org*

Key Words: Teaching, Statistical Education, Examples from the Law

Keeping undergraduate students and medical residents interested in statistics when the majority are taking the class as a requirement can be challenging. I have attempted to accomplish this goal by the use of interesting examples from the law. Examples will include the purported first courtroom case in the US to use statistical evidence, examples of the misuse of statistics that lead to cases being overturned on appeal, a paradox concerning Corporate Average Fuel Economy (CAFE) regulations, and "what went wrong with Executive Order No. 11497."

Consulting Activities and Introductory Course Topics: The Chicken, the Egg, and the Chef?

✱Jim Bentley, University of Redlands, Dept. of Mathematics, 1200 E Colton Avenue, Redlands, CA 92374, *jim_bentley@redlands.edu*

Key Words: undergraduate consulting, introductory course

For many statisticians teaching introductory statistics, the selection of course topics is dictated in part by their understanding of an historic need for certain methods. For many statisticians watching over student consulting experiences, the need to cover topics is mandated by the projects in which the students are involved. In an optimal setting these two sets of topics would be identical. Some would argue that the topics commonly covered in "modern" introductory statistics courses are appropriate for the needs of the students. Experiences in the consulting lab would suggest that we should be looking outside of the classroom to statisticians working within research environments such as industry and government to determine the topics that our students really need to understand.

The 'Parknership' Research Program Between the Florida Park Service and Florida Atlantic University's Honors College

✱Stephanie Fitchett, University of Northern Colorado, Department of Mathematical Sciences, Campus Box 122, Greeley, CO 80639, *stephanie.fitchett@unco.edu*; Hank Smith, Florida Atlantic University & Florida Park Service

Key Words: service learning, statistics education

The "Parknership" Research Program is a wildlife biology research collaboration between the Florida Park Service, Florida Atlantic University, and several other government and educational institutions. Through the "Parknership" program, students and faculty collaborate with scientists in the partner organizations on research related to practical wildlife management. Students who participate in the program are expected to produce work that contributes to effective management of endangered species and their habitats; integrates scientific, political, and economic aspects of wildlife and endangered species management; utilizes appropriate statistical and mathematical tools for data analyses; and is published in scientific journals. This presentation describes the "Parknership" Research Program and the fundamental role of statistics in the program.

101 Nonparametrics and Functional Data Analysis

Section on Statistical Computing, Section on Nonparametric Statistics, IMS

Monday, August 4, 8:30 a.m.–10:20 a.m.

Functional Component Pursuit for Small N, Large P Data

Yoonkyung Lee, The Ohio State University; ✱Yonggang Yao, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210, *yao@stat.osu.edu*

Key Words: Reproducing Kernel, Support Vector Machine, Nonparametric Regularization, Linear Programming

Based on reproducing-kernel Hilbert space theory, a large class of feature spaces can be characterized through their kernel functions, which have been widely applied in modern nonparametric statistics. Nonparametric regularization methods can be used to fit satisfactory predictive models for small n large p data, while their solutions usually involve all the p variables and do not identify the principal components that are more responsible for the prediction. By further parameterizing the kernel structure, many current nonparametric regularization methods can be extended to achieve both model-fitting and feature selection objectives. In this report, we focus on the nonparametric regularization problems associated with a kernel collection, and introduce a fast adaptive kernel selection algorithm rooted in parametric linear programming.

Large-Scale Clustering of Dependent Curves

✱Huijing Jiang, Georgia Institute of Technology, 350448 Georgia Tech Station, Atlanta, GA 30332, *hjiang@isye.gatech.edu*; Nicoleta Serban, Georgia Institute of Technology

Key Words: clustering, spatial dependence, functional data analysis, demographics, Incomplete Choslesky Factorization, kernel-based decomposition

In this paper, we introduce a model-based method for clustering multiple curves or functionals under spatial dependence specified up to a set of unknown parameters. The functionals are decomposed using a semiparametric model where the fixed effects account for the large-scale clustering association and the random effects for the small-scale spatial-dependence variability. The clustering model assumes the clustering membership as a realization from a Markov random field. Within our estimation framework, the emphasis is on a large number of functionals/spatial units with sparsely sampled time points. To overcome the computational cost resulting from large dependence matrix operations, the estimation algorithm includes a two-stage approximation: low-ranked kernel-based decomposition of the dependence matrix and Incomplete Choslesky Factorization of the kernel matrix.

Polynomial Spline Independent Component Analysis with Application to fMRI Data

✱Atsushi Kawaguchi, The University of North Carolina at Chapel Hill, Department of Biostatistic, Chapel Hill, NC 27599, *kawa.a24@gmail.com*; Young K. Truong, The University of North Carolina at Chapel Hill; Xuemei Huang, The University of North Carolina at Chapel Hill

Key Words: Independent Component Analysis, Spline, Spatial-Temporal Data, Functional Magnetic Resonance Imaging

We continue our development of independent component analysis (ICA) using a maximum likelihood approach based on polynomial splines with adaptive knot locations. In a previous study, such approach has produced favorable results based on simulation under the framework of the blind source separation (or temporal ICA), and we also observed that it may not be trivial to apply this method to detect spatial-temporal features from functional magnetic resonance imaging (fMRI) data. In this paper, we incorporate an unsupervised learning method into our algorithm, which yields a significant improvement over existing adaptive methods based on kernel density estimates. The proposed procedure is applied to an fMRI study involving Parkinson's disease patients for identifying regions that are activated by specific experimental tasks.

A Spatial Spline Algorithm and an Application to Climate Waves Over the United States

✱Derek D. Cyr, University at Albany, Dept. of Epidemiology and Biostatistics, School of Public Health, 1 University Place, Rensselaer, NY 12144-3456, *dc896148@albany.edu*; Igor G. Zurbenko, University at Albany

Key Words: Splines, Smoothing, Temperature

Technology of multidimensional signal processing with high levels of noise will be discussed together with a new algorithm of spline smoothing. A 3D spatial KZ spline has been used by I. Zurbenko, R. Henry and S. Sanapati to obtain 4D images of El Niño oscillations in the air temperature over the USA; see *ftp://ftp.dec.state.ny.us/dar/air_research/htdocs* (tmpg folder). These images are constructed as a movie in a compressed time with waves of temperature deviations from normal, scaled in color; it may identify potential disastrous areas in time and space as well as a long term forecast of deviations from normal levels. Such deviations can be predicted using a 3D space approach approximately one year ahead. For more details, see Kolmogorov-Zurbenko Spline Smoothing and Applications. R package version 1.2.0. *http://cran.us.r-project.org.*

Nonparametric Density Deconvolution by Weighted Kernel Estimators

✱Berwin A. Turlach, National University of Singapore, Dept of Statistics and Appl Prob, 6 Science Drive 2, Singapore, 117546 Singapore, *statba@nus.edu.sg*; Martin L. Hazelton, Massey University

Key Words: Density estimation, Errors in variables, Integrated square error, Measurement error, Weights

Nonparametric density estimation in the presence of measurement error is considered. In this paper a new approach based on a weighted kernel density estimator is proposed as opposed to the usual kernel deconvolution

estimator which uses a modified kernel. In practice a data driven method of weight selection is required. Our strategy is to minimize the discrepancy between a standard kernel estimate from the contaminated data on the one hand, and the convolution of the weighted deconvolution estimate with the measurement error density on the other hand. We consider a direct implementation of this approach, in which the weights are optimized subject to sum and non-negativity constraints, and a regularized version in which the objective function includes a ridge-type penalty. Numerical tests suggest that the weighted kernel estimation can lead to tangible improvements.

On the Change-Point Regression by Local PTh-Order Polynomial Kernel Estimators

✱Bo Zhang, University of Minnesota., 1060 27th Ave SE Apt E, School of Statistics, Minneapolis, MN 55414, *dzhang@stat.umn.edu*

Key Words: change-point, nonparametric regression, discontinuity detection, Hausdorff distance

Important issue arises when dealing with nonparametric regression with discontinuities. That is knowledge of discontinuities, including the estimates of not only discontinuities locations but also their magnitudes. In this article, we compare the performance of general estimate $M^p(x)$ ($p=1,2,\cdots$) of absolute difference of two one-sided local pth-order polynomial kernel estimators for potential jump points in conducting detection of discontinuities and estimation of their magnitudes. Theoretical results ensure strong consistency of proposed procedure. The notice is brought to the procedures by one-sided quadratic kernel estimators, and even cubic estimators. The simulation results and real examples show that they outperform the traditional ones in some cases, in sense of Hausdorff distance of estimated jumps and truth. Joint work with Prof Peihua Qiu.

Statistical Inference for Dynamic Models with the Generalized Profiling Method

✱Jiguo Cao, Simon Fraser University, Burnaby, BC V5A1S6 Canada, *jca76@sfu.ca*

Key Words: differential equations, penalized smoothing, profiling method, nuisance parameters, Dynamic models, funcitonal data analysis

Dynamic models, usually written in forms of differential equations (DEs), describe the rate of change of a process. They are widely used in medicine, engineering, ecology and a host of other applications. One central and difficult problem is how to estimate DE parameters from noisy data. We have developed the generalized profiling method to solve this problem. DE solutions are approximated by nonparametric functions, which are estimated by penalized smoothing with DE-defined penalty. The computation is much faster than other methods. A modified delta method is proposed to estimate variances of DE parameters, which include all the uncertainty of the smoothing process. I will demonstrate our method with estimating a predator-prey dynamic model and gene regulatory networks. The generalized profiling method can also be used to estimate other statistical models with nuisance parameters.

102 Recent Developments in Control Charts and Profile Monitoring

Section on Quality and Productivity, Section on Physical and Engineering Sciences

Monday, August 4, 8:30 a.m.–10:20 a.m.

A Distribution-Free, Shewhart-Type Quality Control Chart for Detecting a General Change in a Process

✱Saad T. Bakir, Alabama State University, 2405 Reston Place, Montgomery, AL 36117, *sbakir@alasu.edu*

Key Words: Nonparametric, average run length

A distribution-free (or nonparametric) Shewhart-type quality control chart for detecting a general change in the probability distribution of a process is proposed. The chart is based on the two-sample Kolmogorov-Smirnov statistic and it requires the availability of a reference (or training) sample of observations taken when the process was operating in-control. One advantage over the traditional distribution-based charts is that the proposed chart has the same in-control ARL and false alarm rate over the class of all continuous distributions. Therefore its control limits need not be adjusted according to the assumed underlying distribution. Simulation studies show that the proposed chart can be more efficient than the Shewhart X-bar chart when the process underlying distribution is not normal, especially when this distribution has heavy tails.

Profile Monitoring for Attributes Data

✱Arthur Yeh, Bowling Green State University, Department of Applied Statistics and OR, Bowling Green, OH 43403, *byeh@bgsu.edu*

Key Words: Profile Monitoring, Phase I Control, Attributes Response, Logistic Regression, Covariance Matrix, Shewhart Chart

Pertaining to industrial applications in which the response variable of interest is binary, we study how the profile functional relationship between the response and predictor variables can be monitored using logistic regression. Under such a premise, we extend and study several Hotelling T^2-charts that have been studied under continuous response variable to binary response variable for the purpose of Phase I profile monitoring. Based on simulations, we compare the performance of these T^2-charts charts in terms of the signal probability for different out-of-control scenarios. A real example originated from aircraft construction is given in which these T^2-charts are applied and compared using the data. A discussion of potential future research is also given.

Monitoring Profile Data: A Likelihood Ratio Test Approach

✱Junjia Zhu, The Pennsylvania State University, 331A Thomas Building, University park, PA 16802, *junjia@psu.edu*

Key Words: Statistical Process Control, Profile Monitoring, Generalized Linear Models, Likelihood Ratio Test, EWMA chart, ARL

In this article we focus on Phase II monitoring of profile data. The profile data is assumed to be adequately modeled by the Generalized Linear Model (GLM). Likelihood ratio test (LRT) based control charts are proposed to detect the shift of on-line profiles from the baseline profile. In addition, both Shewhart-type and EWMA-type LRTcontrol charts are considered.

Diagnostics After a Signal from Control Charts in a Normal Process

✱Jianying Lou, Virginia Polytechnic Institute and State University, Department of Statistics, Virginia Tech, Blacksburg, VA 24061, *jlou@vt.edu*; Marion R. Reynolds, Jr., Virginia Polytechnic Institute and State University; Dong-Yun Kim, Virginia Polytechnic Institute and State University

Key Words: Control chart, signal, diagnostics, maximum likelihood

Control charts are fundamental SPC tools for process monitoring. When a combination of charts signals, knowing the change point, which distributional parameter changed, and/or the change size helps to identify the cause of the change. Traditionally the change point and the changed parameter are determined by informally examining the control charts. Maximum likelihood (ML) estimation of the change point is also used in SPC, but estimating the current process mean associated with special causes has mainly been done in EPC. In this paper, we develop diagnostics to provide information about the change point, the changed parameter and the change size. We propose using ML estimators of the current process parameters and their confidence intervals to identify and estimate the changed parameters. This approach works well in most cases, and has better performance than the traditional approach.

CUSUM Charts with Censored Ordinal Categorical Data

✱Denisa A. Olteanu, Virginia Polytechnic Institute and State University, 615 Washington ST SE, Apt 3, Dept. of Statistics, Blacksburg, VA 24060, *dolteanu@vt.edu*

Key Words: Ordinal Categorical, Censored, quality control, CUSUM

In many quality control applications the data collected fall into a finite set of ordered categories. In some cases the analyst only knows an upper or lower limit for the actual data point. This paper uses a latent continuous random variable that takes values in the intervals determined by the ordered categories to model the process. The proposed procedure uses samples of n data points and models the counts of data points falling in any specific interval by a multinomial distribution. This paper develops a cumulative sum (CUSUM) control chart based on the multinomial distribution to detect changes in the latent distribution. This paper illustrates the proposed procedure to one of Becton Dickinson's, the medical devices manufacturer, processes.

Control Charts with Missing Observations

✱Sara R. Wilson, Virginia Polytechnic Institute and State University, 307 Hunt Club Road #6400 I, Blacksburg, VA 24060, *sawilson@vt.edu*; Marion R. Reynolds, Jr., Virginia Polytechnic Institute and State University

Key Words: Statistical process control, Control charts, EWMA, MEWMA, Average time to signal

Traditional control charts for process monitoring are based on taking samples from the process at regular time intervals. However, it is often possible in practice for observations, and even entire samples, to be missing. We investigate different ways to deal with missing observations in Exponentially Weighted Moving Average (EWMA) control charts by considering several approaches for adjusting the weights of the EWMA control statistic. The case in which individual observations are taken at each sampling point, as well as the case in which the sample size is greater than one are considered. Integral equation and Markov chain methods are developed to find and compare the statistical properties of these charts. The multivariate case in which information on some of the variables is missing is also examined using Multivariate EWMA (MEWMA) control charts.

A Residual-Based Hotelling's T^2 Chart

Soren Bisgaard , University of Massachusetts Amherst; ✱Xuan Huang, University of Massachusetts Amherst, Isenberg School of Management, 121 Presidents Drive, Amherst, MA 01003, *xuan@som.umass.edu*

Key Words: Process monitoring, Statistical process control, Quality control, Multivariate time series, Average run length

The application of Hotelling's T^2 for the simultaneous monitoring of several quality characteristics assumes temporal independence. However, industrial data is often autocorrelated. For applications where the data stream follows a multivariate time series model, we propose a special cause control chart based on applying Hotelling's T^2 to the residuals. We provide detailed evaluations of the performance of the proposed method based on the Average Run Length (ARL) for a group of first order vector autoregressive models. We conclude that the residual based method in most cases perform favorably relative to using a Hotelling's T^2 chart directly on the unfiltered observations. We also demonstrate the method when applied to data from an industrial process.

103 Nonparametric Identification of Densities, Features, Mixtures, Shapes, and Images

Section on Nonparametric Statistics, IMS
Monday, August 4, 8:30 a.m.–10:20 a.m.

Analysis of the Shape of Unimodal Densities with Nonparametric Density Estimation

✱Jussi Klemela, University of Oulu, Department of Mathematical Sciences, P.O. Box 3000, Oulun Yliopisto, International 90014 Finland, *klemela@oulu.fi*

Key Words: multivariate data, nonparametric density estimation, level set trees, shape analysis, visualization

Nonparametric density estimation has been applied successfully in mode detection. However, nonparametric analysis of the shape of unimodal densities has attracted less interest. Level set tree based techniques can be applied to analyze the shape of unimodal densities. A level set tree of a function is a tree structure of the separated components of level sets of the function. Level set trees can be used to describe not only the shape of functions but also the shape of multidimensional sets; we can define a distance function or a height function on a set and construct a level set tree of this function. Finally, level set trees can be used to describe the shape of point clouds, by applying appropriate smoothing. This leads to a computationally efficient way of describing the shape of unimodal densities.

A Nonparametric Kurtosis Ordering for Multivariate Distributions

✱Jin Wang, Northern Arizona University, Department of Mathematics and Statistics, Flagstaff, AZ 86011-5717, *jin.wang@nau.edu*; Weihua Zhou, The University of North Carolina, Charlotte

Key Words: Nonparametric method, Volume functional, Kurtosis, Ordering, Elliptically symmetric distributions, Multivariate normality

Extending the Balanda and MacGillivray (1990) kurtosis ordering for univariate distributions, we define and study a nonparametric kurtosis ordering for multivariate distributions. This ordering is preserved by the nonparametric multivariate kurtosis functional of Wang and Serfling (2005) and it is inversely preserved by the "fan plot" of Liu, Parelius and Singh (1999). For elliptically symmetric distributions, the ordering determines the distribution up to affine equivalence. This is applied to design a new graphical method to assess multivariate normality. Ordering results are established for three important families of elliptically symmetric distributions: Kotz-type distributions, Pearson Type VII distributions, and Pearson Type II distributions.

Nonparametric Estimation in Multivariate Mixture Models

✱Tatiana Benaglia, The Pennsylvania State University, 325 Thomas Building, University Park, PA 16802, *tab321@stat.psu.edu*; David Hunter, The Pennsyvania State University; Didier Chauveau, Université d'Orléans, France

Key Words: EM algorithm, kernel density estimation, multivariate mixture, nonparametric mixture

We propose an algorithm for nonparametric estimation for finite mixtures of multivariate random vectors that is not, but strongly resembles, a true EM algorithm. The vectors are assumed to have conditionally independent coordinates, but otherwise their density functions are completely unspecified. Our algorithm is much more flexible and easily applicable than existing algorithms in the literature; it can be extended to any number of mixture components and any number of coordinates of the multivariate observations. Thus it may be applied even in situations where the model is not identifiable, so care is called when it is difficult to establish identifiability conclusively. Our algorithm yields much smaller mean integrated squared errors than an alternative algorithm in a simulation study. In another example using a real data set, it provides new insights that extend previous analyses.

L2 Estimation of Mixture Complexity

✱Umashanger Thayasivam, The University of Georgia, Department of statistics, 101 Cedar Street, Athens, GA 30602-1952, *rtshanger@yahoo.com*; T. N. Sriram, The University of Georgia

Key Words: Mixture Complexity, Minimum Integrated Square distances, robust estimation, minimum distance

Developing statistical procedures to determine the number of components, known as mixture complexity, remains an area of intense research. In many applications, it is important to find the mixture with fewest components that provides a satisfactory fit to the data. Here, we focus on consistent estimation of unknown number of components in finite mixture models, when the exact form of the component densities are unknown but are postulated to be close to members of some parametric family. Minimum Integrated Square distances (L2E) are used to develop a robust estimator of mixture complexity, when all the parameters associated with the model are unknown. The estimator is shown to be consistent. We illustrate the use of our method for three well known datasets: the acidity data, enzyme data and galaxy data.

Probabilistic Watermark Detection in Movies

Sam Behseta, California State University, Bakersfield; Charles Lam, California State University, Bakersfield; ✱Robert Webb, California Polytechnic State University, San Luis Obispo, CA 93407, *webb@calpoly.edu*

Key Words: Image Processing, Nonparametric Bootstrap, Watermarking, Discrete Cosine Transforms

In this work, we propose different algorithms to approximate an original watermarked document, primarily a movie represented with a sequence of matrices. We consider the case of a collusion attack when a number of watermarked copies are available. We use the Bootstrap to construct point-wise confidence intervals for each matrix, represented via their discrete cosine transforms. We demonstrate, by way of an extensive simulation that in addition to stellar probabilistic coverage, Bootstrap machinery is highly efficient when the number of Bootstrap iterations, the sample size and the number of watermarked copies are of interest. Most importantly, the results suggest that the precision of our probabilistic methodology increases quickly when the number of watermarked copies are increased.

Density Estimation Through Kernel Estimation-Based ECF

✱Mawia Kaddoura, The Petroleum Institute, United Arab Emirates, *mkaddoura@pi.ac.ae*; Michael Conerly, The University of Alabama

Key Words: kernel function, characteristic function, ECF, evaluation of bandwidth

This paper reports on an investigation of the use kernel methods to estimate the CF. The proposed method enhances the smoothing of the ECF to improve its performance and is based on using the CF of the kernel method to estimate the underlying distribution without having to calculate the optimal value of the kernel smoothing parameter. Smoothing the ECF by multiplying it by the CF of the kernel is equivalent to the estimation of the CF via the kernel method. The proposed method calls for minimizing the l2 norm of the mean squared error of the estimated CF using the kernel method with respect to the bandwidth, which yields an estimator that does not depend directly on any kernel function and avoids the evaluation of bandwidth. The inversion theorem of the resulting CF is used to estimate of the underlying PDF. Supporting simulation results will be included in the paper.

104 Model Selection, Matching, and Meta-Analysis

Section on Health Policy Statistics, Biopharmaceutical Section, Section on Statistics in Epidemiology

Monday, August 4, 8:30 a.m.–10:20 a.m.

Use of Propensities with Unmeasured Confounders

✱David Nelson, Minneapolis VA Medical Center, University of Minnesota, One Veterans Drive, Minneapolis, MN 55417, *dave.nelson@med.va.gov*; Siamak Noorbaloochi, VAMC and University of Minnesota

Key Words: Propensity Theory, Missing Confounders, Dimension Reduction

Observational studies assessing causal relationships between an explanatory measure and an outcome can be complicated by hosts of potential confounding measures. Propensity theory effectively addresses confounding when all confounders are measured. The assumptions underlying nonconfounded causal inference using propensities breakdown when there are unmeasured confounders. However, when there are unmeasured confounders, the dimension reduction central to propensities and related summaries can still be utilized in the inferential process. We discuss how propensity theory can be combined with simple covariate free methods for addressing unmeasured confounders to estimate the causal effects.

Bayesian Models for the Meta-Analysis of Sparse Tables

✱Eloise Kaizar, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210, *ekaizar@stat.osu.edu*; Joel Greenhouse, Carnegie Mellon University; Howard Seltman, Carnegie Mellon University

Key Words: meta-analysis, Bayesian models, rare events, sparse tables

Recent concern around the possible association of severe but rare adverse events with commonly-used medications (e.g., antidepressants, COX-2 inhibitors) has led to a number of meta-analyses of adverse events observed in randomized controlled trials. Because so few events are observed in each trial, many traditional meta-analysis methods are computationally unable to be applied in these analyses without ad-hoc adjustment. Bayesian models have been proposed to overcome many of the computational issues associated with rare event meta-analysis. However, the behavior of these models has not been adequately explored. We illustrate how different parameterizations and corresponding choice of prior specification impact the performance of Bayesian estimates and their robustness.

The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation

✱Kosuke Imai, Princeton University, Department of Politics, Princeton, NJ 08544, *kimai@Princeton.Edu*; Gary King, Harvard University; Clayton Nall, Harvard University

Key Words: causal inference, community intervention trials, group-randomized trials, matched-pair design, health policy, noncompliance

A basic feature of many field experiments is that investigators are only able to randomize clusters of individuals even when individuals are the unit of interest. To recoup some of the resulting efficiency loss, many studies use the "matched-pair cluster-randomized design." Other studies avoid pairing, because some claim to have identified serious problems with this design. We prove that all such claims are unfounded, and show that the estimator favored in the literature is appropriate only in situations where matching is not needed. To address this problem, we propose a simple nonparametric estimator with improved statistical properties. We show that from the perspective of bias, efficiency, power, or robustness, pairing should be used whenever feasible. We develop these techniques in the context of a randomized evaluation of the Mexican Universal Health Insurance Program.

Model Selection for Individualized Treatment Rules

✱Min Qian, The University of Michigan, 439 West Hall, 1085 South University Ave., Ann Arbor, MI 48104, *minqian@umich.edu*; Susan A. Murphy, The University of Michigan

Key Words: decision making, machine learning, oracle inequality

In medicine, no single treatment has been shown to be effective for all. This has motivated researchers to incorporate the heterogeneity in need for treatment across individuals. A decision rule that assigns treatment according to patient characteristics is called an individualized treatment rule. Our goal is to estimate the treatment rule that maximizes the mean reward. Penalized regression is used to estimate the optimal individualized treatment rule. We first give nontrivial upper bounds on the excess mean reward in terms of the excess prediction error. Fast rate of convergence of the excess mean reward is obtained in the case of low noise. Then we provide a non-asymptotic upper bound of the excess prediction error. Finally we present an oracle inequality of the excess mean reward.

Model Comparison for Association Between Hospital Performance and Hospital Characteristics: Proportional or Hierarchical Binary Outcomes

✱Jacob J.H. Cheng, Maryland Hospital Association, 6816 Deerpath Rd, Quality Indicator Project, Elkridge, MD 21075, *jcheng@mhaonline.org*; Carlos F. Alzola, Data Insights; Nikolas Matthes, Maryland Hospital Association

Key Words: Proportional Outcome, Hierarchical Binary Outcome, Linear Models, Generalized Linear Mixed Models (GLMM), Generalized Estimating Equations (GEE), Hospital Performance and Characteristics

Hospital performance is usually measured by the proportion of patients who received appropriate care. Linear models are intuitive tools to study the association between performance and hospital characteristics but may violate normality and homoscedasticity assumptions. GEE and GLMM work on the binary patient-level outcomes. 2005–2006 data for heart failure care from 442 QIP hospitals were used. Models were evaluated by R2/RMSE with cross-validation and Type III test. Arcsine transformation improved the above problem in linear models resulting in closer inference to GEE and GLMM. GEE had the best power for marginal predictions while GLMM provided additional hospital-specific predictions. The difference between linear models and GEE/GLMM was minimal with enough sample size, leaving linear models attractive based on their flexibility for handling model diagnosis and nonlinearity detection.

Longitudinal Data with Follow-Up Truncated by Death: Communicating a Match Between Analysis Method and Research Aims

✱Laura L. Johnson, NCCAM/NIH, 6707 Democracy Blvd, Suite 401, MSC 5475, Bethesda, MD 20892-5475, *johnslau@mail.nih.gov*; Brenda Kurland, Fred Hutchinson Cancer Research Center; Paula Diehr, University of Washington

Key Words: longitudinal data, truncation by death, missing data, generalized estimating equations, random effects models

How to summarize trajectories of longitudinal data truncated by death not non-response? Unconditional models such as random effects models may implicitly impute data beyond the time of death. Fully conditional models stratify longitudinal response trajectories by time of death. Partly conditional models reflect the average response in survivors at a given time point, rather than individual trajectories. Joint models of survival and longitudinal response describe the evolving health status of the entire cohort. Researchers using longitudinal data should consider which method of accommodating deaths is consistent with research aims, and use analysis methods accordingly. While subtly different wording is used to describe the analyses, the results can vary dramatically, impacting clinical and public interpretation of an analysis.

105 Modeling, Estimation, and Testing for Stochastic Processes

IMS

Monday, August 4, 8:30 a.m.–10:20 a.m.

Statistical Tests for the Detection of a Self-Regulation Property for Poisson-Type Point Processes

✱Franz Streit, University of Geneva, Department of Mathematics, Case Postale 64, CH-1211 Geneva 4, CH 1211 Switzerland, *streit@math.unige.ch*

Key Words: Locally most powerful tests, Asymptotic optimality of tests, Point processes, Self-regulation property

For an orderly point process on the positive real numbers let t(i) be the ith point event after t=0 and the complete intensity function be given by g(t:r;d,c)=d+r1(t-t(i-1),{t-t(i-1)>c}) for t in (t(i-1),t(i)], i=1,2,...; d,c>0; r nonnegative; t(0)=0 and 1(x,A)=1 if x satisfies the condition A and 1(x,A)=0 otherwise. We present tests for the hypotheses H: r=0 and K: r>0 for known d and c based on the observation of the point process in the time period (0,t(m)], where m is a fixed natural number. We indicate a locally most powerful test and look then at the problem from the point of view of asymptotic optimality of the procedure (contiguous hypotheses). These results are extended to other self-regulation properties.

Finding the Recurrent Sets for a Markov Chain Possessing Monotonically Decreasing Nonhomogeneity

✱Zach Dietz, Hamilton College, 20 College Hill Rd, Apartment 7, Clinton, 15213, *zdietz@hamilton.edu*

Key Words: Markov chain, non-homogeneity, location statistic

Consider a finite state, non-homogeneous Markov chain with transition kernels \{P_n\}, where P_n monotonically converges to P in a pointwise fashion. It has long been known that the limiting distribution of X_n, the location statistic of the chain at time n, gives probability 1 to a set of recurrent states. We present some partial results appertaining to what this limiting distribution is, an algorithm used to determine the recurrent states of this limiting distribution.

Multivariate Stationary Processes with Increments

✱Chunfeng Huang, Indiana University, Statistics House, 309 North Park Avenue, Bloomington, IN 47408, *huang48@indiana.edu*; Tailen Hsing, The University of Michigan

Key Words: spectral representation, correlation theory, random field, stochastic process

To study the correlation of a multivariate stationary process with increments which may possess different orders, a spectral representation theorem is obtained for the correlation functions and the process. Some examples and applications are presented. This general theorem covers the result in Yaglom (1958) and can be extended to the study of random fields.

Finite-Sample Properties of a System Reliability Estimator Under Imperfect Debugging

✱Marcus Agustin, Southern Illinois University Edwardsville, Department of Mathematics and Statistics, Edwardsville, IL 62026, *magusti@siue.edu*; Zenia Agustin, Southern Illinois University Edwardsville

Key Words: counting processes, software reliability

We will consider a system of components where the failure of each component depends on the number of defects in the system and the failure rate contribution of each defect. The number of defects changes dynamically since removal of a defect may result in new defects being introduced to the system. The system is observed for a fixed time period and the main interest is to obtain an estimator of the system reliability at the end of the observation period. The finite-sample properties of the resulting estimator will be evaluated using a simulation study.

Nonparametric Covariogram Estimation Based on Irregularly Spaced Spatial Data

✱Soutir Bandyopadhyay, Texas A&M University, 3143 TAMU, Department of Statistics, College Station, TX 77843-3143, *soutir@stat.tamu.edu*; Soumendra N. Lahiri, Texas A&M University

Key Words: spectral density, covariogram, spatial process, stochastic design

We propose a nonparametric estimator for the covariogram function based on irregularly spaced spatial data generated by a stochastic design. Unlike earlier works on the problem, the proposed estimator of the covariogram does not involve any smoothing. As a result, the proposed method avoids the vexing problem of finite sample choice of the smoothing parameter for practical applications. We investigate theoretical properties of the covariogram estimator under pure- and mixed-increasing domain spatial asymptotics. We will also present results from a moderately large simulation study.

Nonexplosion of a Class of Semilinear Partial Differential Equations via Branching Particle Representations

✱Santanu Chakraborty, University of Texas - Pan American, 1201 West University Drive, Edinburg, TX 78541, *schakraborty@utpa.edu*; Jose A. Lopez Mimbela, Centro de Investigacion en Matematicas

Key Words: Branching particle systems, Semilinear partial differential equations, Global positive solutions

We consider a branching particle system where an individual particle gives birth to a random number of offspring at the place where it dies before its death. We assume a discrete probability distribution on the number of offspring. The corresponding branching process is described by a semilinear partial differential equation. We obtain sufficient conditions for existence of global positive solutions to semilinear equations of this form. These conditions answer questions regarding a population becoming extinct in future.

106 Bayesian Methods; Ridge Regression

Business and Economics Statistics Section, Section on Bayesian Statistical Science

Monday, August 4, 8:30 a.m.–10:20 a.m.

A Bayesian Approach to Nonparametric Monotone Function Estimation

✱Tom Shively, University of Texas at Austin, IROM Department, Mail Code B6500, University of Texas, Austin, TX 78712, *tom.shively@mccombs.utexas.edu*; Tom Sager, University of Texas at Austin

Key Words: MCMC sampling scheme, Mixture prior distributions, Regression splines, Small sample properties

Two Bayesian approaches to nonparametric monotone function estimation are proposed. The first approach uses a hierarchical Bayes framework and a characterization of smooth monotone functions given by Ramsay (1998) that allows unconstrained estimation. The second approach uses a Bayesian regression spline model (Smith and Kohn, 1996) with a mixture distribution of constrained normals as the prior for the regression coefficients to ensure the monotonicity of the resulting function estimate. The small sample properties of the two function estimators across a range of functions are provided via simulation and compared with existing methods. An example is provided involving economic demand functions that illustrates the application of the constrained regression spline estimator in the context of a multiple regression model where two functions are constrained to be monotone.

Score-Based Portfolio Strategy in Bayesian View

✱Timothy H. Lee, CRM, 595 Croydon Lane, Alpharetta, GA 30022, *tlee080551@aol.com*

Key Words: Score Strategy, Portfolio decision, Bayesian

In some portfolio decision making, a target population is split into subgroups based on selected criteria then various available actions are assigned to the subgroups. Such a practice is usually based on statistically derived scores using two group classification method. The scores are sometimes combined with other score(s) or predictive variable(s)in a matrix or in an equation and then performance for each combination (cell)or score group is observed for portfolio decision. In this study, commonly practiced score strategy was discussed in a Bayesian view point and a combination of an intuitive business approach and a statistical method was suggested for portfolio management in practitioners' point of view.

Bayesian Inference in the Time-Varying Cointegration Model

Gary Koop, University of Strathclyde; ✱Roberto Leon-Gonzalez, GRIPS, 7-22-1, Roppongi, Minato-Ku, Tokyo, 106-8677 Japan, *rlg@grips.ac.jp*; Rodney Strachan, University of Strathclyde

Key Words: Bayesian, directional statistics, error correction model, reduced rank regression, Markov Chain Monte Carlo, time varying cointegration

There are both theoretical and empirical reasons for believing that the parameters of macroeconomic models may vary over time. However, work with time-varying parameter models has largely involved Vector autoregressions (VARs), ignoring cointegration. In this paper we develop time varying parameter models which permit cointegration. Time-varying parameter VARs (TVP-VARs) typically use state space representations to model the evolution of parameters. In this paper, we show that it is not sensible to use straightforward extensions of TVP-VARs when allowing for cointegration. Instead we develop a specification which allows for the cointegrating space to evolve over time in a manner comparable to the random walk variation used with TVP-VARs. The properties of our approach are investigated before developing a method of posterior simulation. An empirical illustration is provided.

Generalized Maximum Entropy Estimation of the Parameters of Production and Demand Functions in the Framework of CGE Models

✱Tiziana Laureti, University of Tuscia, Via del Paradiso 47, Department of Economics, Viterbo, 01100 Italy, *laureti@unitus.it*; Guido Ferrari, University of Florence; Luca Secondi, University of Florence

Key Words: CGE models, RESAM, ill-posed, GME, production functions, demand functions

In the computation of a General Equilibrium Model (CGE), based on a Social Accounting Matrix (SAM), the estimation of behavioral parameters plays an important role. Usually, these parameters are estimated using time series or cross-section data or, when additional information are not available, they are taken from similar contexts in literature. However, if the data used do not reflect the same level of aggregation, comparability and classification problems arise. The approach proposed in this paper is based on the macroeconomic information contained in a Regional Environmental SAM (RESAM) for an Italian region in order to estimate the production and demand functions' parameters. As a RESAM does not contain enough information to obtain statistically significant estimates, we will make resort to the Generalized Maximum Entropy (GME) method, which provides a flexible estimation framework.

Stochastic Volatility Model with Regime Switching

✱Lu Zhang, The Pennsylvania State University, 325 Thomas Building, PSU, University Park, PA 16802, *lzz114@psu.edu*; John Liechty, The Pennsylvania State University

Key Words: Regime switching stochastic volatility, Monte Carlo Markov Chain

We model the unobserved regimes as a continuous-time hidden Markov chain, with exponential waiting times. This allows us to identify the hidden regimes and compare changes in regimes with news events. Our model is very flexible: the number of possible regimes, the mean reverting, persistence and variance parameters in the volatility equation can all be controlled by different hidden Markov chains. We use Monte Carlo Markov Chain method for inference; model selection is based on calculating posterior probability of the model, given the data via a reversible jump approach. We also consider maximum likelihood, and out of sample predictive validation. We report the analysis of several empirical data sets—S&P 500, federal rate, exchange rate—and demonstrate proposed models outperform benchmark models in fit and predictive ability.

107 Longitudinal Data Analysis I

Biometrics Section, Biopharmaceutical Section, Section on Statistics in Epidemiology, Section on Survey Research Methods

Monday, August 4, 8:30 a.m.–10:20 a.m.

Canonical Correlation Analysis of Longitudinal Data

✱Dayanand N. Naik, Old Dominion University, Department of Mathematics and Statistics, Norfolk, VA 23529, *dnaik@odu.edu*; Jayesh Srivastava, High Point Insurance

Key Words: Canonical correlations, Repeated Measures

Studying the relationship between two multivariate random vectors is an important problem in statistics. Canonical correlation coefficients are used to study these relationships. Canonical correlation analysis (CCA) is a general multivariate method that is used to study the relationship when both sets of variables are quantitative. In this talk, we provide a generalization of this method to determine the relationships between two sets of repeatedly or longitudinally observed vectors. Assuming a block Kronecker product variance covariance matrix to account for the dependency of the vectors observed over t time periods, we provide methods to obtain canonical correlations and canonical variables.

Inverse Regression from Longitudinal Data

✱Geoffrey Jones, Massey University, Private Bag 11222, Palmerston North, New Zealand, *g.jones@massey.ac.nz*

Key Words: Calibration, Hierarchical model, Multilevel model, MCMC

Inverse regression, or statistical calibration, uses the estimated relationship between a response Y and a covariate x to infer the values of unknown x's from their observed Y's. Typically x is univariate but Y may be multivariate. A brief review of the basic theory will be given, followed by consideration of the problems involved in extending these approaches to longitudinal data, ie where the training data consists of groups of observations on distinct individuals. A Bayesian analysis using MCMC is shown to give a flexible framework for solving these problems. An example concerning the age determination of tern chicks from their wingspan and weight measurements will be used for illustration.

State Space Representation of an Autoregressive Linear Mixed Effects Model for the Analysis of Longitudinal Data

✱Ikuko Funatogawa, Teikyo University School of Medicine, 2-11-1 Kaga, Itabashi-ku, Tokyo, 173-8605 Japan, *funatogawaikkjp@yahoo.co.jp*; Takashi Funatogawa, Chugai Pharmaceutical Co. LTD.

Key Words: Autoregressive model, Equilibrium, Kalman filter, Linear Mixed Effects Model, Longitudinal data, State Space

Recently, we proposed an autoregressive linear mixed effects model for the analysis of longitudinal data in which the current response is regressed on the previous response, fixed effects, and random effects (Funatogawa et al. 2007 Stat Med). The model represents profiles approaching random equilibriums. Because intermittent missing is an inherent problem of the autoregressive (conditional) mode, we provided the marginal (unconditional) representation of the model and the likelihood. In this study, we further provide a state space form of our model for calculating the marginal likelihood without using large matrices. We modified the method proposed by Jones (1993) for a state space form of a usual linear mixed effects model. We analyzed parathyroid hormone and serum calcium measurements in treatment of secondary hyperparathyroidism.

Estimation of Disease Progression in Amyotrophic Lateral Sclerosis Using Muscle Electrical Properties

✱Catherine Stamoulis, Harvard School of Public Health, 655 Huntington Ave, Department of Biostatistics, Building 2, Room 428, Boston, MA 02115, *caterina@mit.edu*; Rebecca Betensky, Harvard University; Seward Rutkove, Harvard Medical School

Key Words: Amytrophic Lateral Sclerosis, progression, mixed-effects models, biomarkers, sampling

Amyotrophic Lateral Sclerosis (ALS) is a fatal neurodegenerative disease characterized by progressive muscle weakness, for which only relatively insensitive markers exist. ALS varies highly among people, with non-linear initial progression followed by linear decline. Electrical impedance myography was used to measure muscle impedance longitudinally in 18 ALS patients at different disease stages, 27 normal subjects and 20 patients with disuse myopathy. The data were fitted with a mixed effects model to account for disease duration, time between measurements and covariation of measurements from muscles in the same anatomical group. Sensitivity and specificity of multi-frequency resistance, reactance and impedance phase were assessed for progression. Electrical resistance differentially varied in ALS patients; reactance and phase spectra varied non-specifically in ALS and disuse myopathy.

Comparison Between Analysis of Variance and Spectral Decomposition in the Mixed-Effects Models

✱Mixia Wu, National Institute of Child Health and Human Development, 6100 Executive Boulevard, Rm 7B05A, Rockville, MD 20852, *wumixia@mail.nih.gov*; Kai Fun Yu, National Institutes of Health; Aiyi Liu, National Institute of Child Health and Human Development

Key Words: Analysis of variance, Best linear unbiased estimate, Spectral decomposition, Mixed-effects model, Least squares estimate

The mixed-effects models with two variance components are often used to analyze longitudinal data. For this kind of model, we compare the two approaches to estimating the variance components, the analysis of variance approach and the spectral decomposition approach. We establish a necessary and sufficient condition for the two approaches to yield identical estimates, and some sufficient conditions for the superiority of one approach over the other, under the mean squared error criterion. Applications of the methods to circular models and panel data are discussed. Furthermore, simulation results indicate that better estimates of variance components do not necessarily imply higher power of the tests or shorter confidence intervals.

Marginalized Transition Models for Longitudinal Count Data

✱Keunbaik Lee, Louisiana State University Health Sciences Center, 2021 Lakeshore Drive, Suite 210, New Orleans, LA 70112, *klee4@lsuhsc.edu*; Michael Daniels, University of Florida; Yongsung Joo, University of Florida

Key Words: poisson, overdispersion, Fisher-scoring

Heagerty (2002) and Lee and Daniels (2007) have proposed marginalized transition models for the analysis of longitudinal binary data and ordinal data, respectively. In this paper, we propose similar models for longitudinal count data. We also propose a model to accommodate overdispersed count data. Fisher-scoring algorithms are developed for estimation. Methods are illustrated with a real dataset and are compared with other standard methods.

108 Pathway Analysis Methods in Genomics

Biometrics Section, IMS

Monday, August 4, 8:30 a.m.–10:20 a.m.

Exploring Liquid Association in Non-Gaussian Multivariate Distributions

✱Yen-Yi Ho, Johns Hopkins University, 500 W. University Parkway, Apt 9S, Baltimore, MD 21210, *yho@jhsph.edu*; Leslie Cope, Johns Hopkins University; Giovanni Parmigiani, Johns Hopkins University

Key Words: Liquid Association, Higher-order interaction, Copula, Gram-Charlier expansion, Estimating Equations

The analysis of higher-order interactions among a group of genes is fundamental to further our understanding of gene regulatory networks. Liquid Association (LA) is an exploratory approach to characterize higher-order interactions among three or more genes. Analysis of gene expression data has suggested that these interactions are present and measurable in many biological systems. However, commonly used multivariate models, most notably the multivariate Gaussian, are limited because the three product-moment including LA are identically 0. In this paper, we examine several statistical models that would allow for tri-variate distributions to display various ranges of LA. These models include copula models, the Gram-Charlier expansion, and the conditional normal model. We provide methods to estimate LA using the models, illustrate these methods in simulation and discuss computation issues.

Concordant Association with Multiple Phenotypes Analysis of Microarray Gene Expression Data

✱Stanley B. Pounds, St. Jude Children's Research Hospital, Dept. of Biostatistics, 332 N. Lauderdale St., Memphis, TN 38105-2794, *stanley.pounds@stjude.org*

Key Words: microarray, permutation, multiple phenotypes, gene expression, biological model

Concordant association with multiple phenotypes (CAMP) analysis is a method that identifies genes showing a pattern of association with multiple phenotypes that is consistent with a biological model. A model of curing leukemia involves (1) increasing drug-concentrations to (2) reduce the rate of DNA synthesis and thus (3) reduce the leukemia burden. In this example, the expression of genes with a "curative" pattern positively correlates with (1), and negatively correlates with (2) and (3). Instead of examining the association with each phenotype separately, CAMP analysis defines a test statistic that measures the evidence for a pattern of association with these variables as being "curative" or "anti-curative" in direction. Significance is determined by permuting the assignment of phenotype data to expression data. CAMP improves power to identify genes with such patterns.

Combining P-Values on Testing the Significance of Groups of Genes

✱Hongmei Jiang, Northwestern University, Department of Statistics, 2006 Sheridan Road, Evanston, IL 60208, *hongmei@northwestern.edu*

Key Words: microarray, gene-set analysis, functional category, Fisher's method of combining p-values

Identifying differential gene expression in microarray studies have recently shifted from testing individual genes to groups of genes, which are defined by biological processes, cellular components, or molecular functions, such as gene ontology (GO) annotations. A list of differentially expressed GO terms is more meaningful to answer biological questions than a list of statistically significant genes. This paper investigates Fisher's method of combining p-values from the gene-basis tests for a group of genes. Different ways to incorporate the correlation among the genes are compared by simulation studies. Other existing methods to combine information across the genes to evaluate the statistical significance of a group of genes are also discussed.

Dynamic Weighted Clustering with Noise Set

✱Yijing Shen, University of California, Los Angeles, 11121 Queensland St, H42, Los Angeles, CA 90034, *yshen@stat.ucla.edu*; Wei Sun, The University of North Carolina at Chapel Hill; Kerchau Li, University of California, Los Angeles

Key Words: clustering, bioinformatics, gene expression, coexpression, gene ontology, EM algorithm

Various clustering methods have been applied to microarray gene expression data to identify genes with similar expression profiles. Functionally related genes may be regulated by common cellular signals, and hence be co-expressed. The biological annotation data have been accumulated and organized into functional categories such as Gene Ontology. How to utilize these rapidly increasing functional annotation resources to improve the performance of clustering methods is of great interest. Forcing scattered genes into existing clusters would reduce the accuracy of the clustering result. In this study, we propose a new approach that incorporates the functional annotation of genes and allows scattered genes to remain unclustered to achieve the goal of function prediction for un-annotated genes.

Identify Relevant Genes for Classification of Microarray Samples Using Partial Least Squares

✱Ying Chen, University of California, Davis, Department of Statistics, Davis, CA 95616, *lynchen@ucdavis.edu*

Key Words: partial least squares, variable selection, dimension reduction, microarray classification, resampling

Partial Least Squares (PLS) has been used as a dimension reduction technique in the context of classification with microarray data and is well known to have excellent predictive accuracy. In order to understand the biology behind the classification models, biologists often want to know which genes contribute most to the variations in the classes of microarray samples. There have been some variable selection methods for PLS, which can give a relative ranking of each gene variable's contribution to the model. However, there is no clear cutoff point. We developed a resampling-based procedure in the statistical testing framework to identify relevant gene variables in PLS models. We tested the procedure in several real microarray data sets and also performed simulation studies. Our proposed method showed superior or comparable performance to some existing methods in different situations.

A Hidden Markov Model Approach to Testing Multiple Hypotheses on a Directed Acyclic Graph

✱Kun Liang, Iowa State University, 428 S Walnut Ave #24, Ames, IA 50010, *liangkun@iastate.edu*; Dan Nettleton, Iowa State University

Key Words: Bayesian data analysis, Multiple testing, Gene enrichment analysis, Gene Ontology, Microarray, Differential expression

Gene category testing problems involve testing hundreds of null hypotheses that correspond to nodes in a directed acyclic graph. The logical relationships among the nodes in the graph imply that only some configurations of true and false null hypotheses are possible and that a test for a given node should depend on data from neighboring nodes. We develop an approach based on a hidden Markov model that takes the whole graph into account and provides coherent decisions in this structured multiple hypothesis testing problem. The method will be illustrated by testing Gene Ontology terms for evidence of differential expression.

109 Statistical Inference in Clinical Trials ●

Biopharmaceutical Section, Biometrics Section
Monday, August 4, 8:30 a.m.–10:20 a.m.

Testing the Assumption in Several Two-Sided Amalgamation-Based Tests for Dose Response

✱Arthur Roth, Pfizer, Inc., 48105, *arthur.j.roth@pfizer.com*

Key Words: dose response, monotonicity, isotonic regression, amalgamation, trend tests, order restriction

Before performing any statistical test that makes an assumption, it is often desirable to do a preliminary test to check the assumption. Several known tests for dose response employ isotonic regression and amalgamated means, including the chibar square test (nonparametric), the ebar square test (parametric with equal variances), and the Brown-Forsythe trend test (parametric with unequal variances, Roth 1983); all of them assume that the means are a monotonic function of dose. For one-sided alternatives, each of them has a known corresponding test for assessing the validity of the monotonicity assumption (one due to Roth 2005) that has no obvious two-sided analogue. We modify all three of the above one-sided tests to obtain appropriate two-sided tests for monotonicity in each setting (nonparametric, parametric with equal variances, and parametric with unequal variances).

Preference-Based Estimates of Treatment Efficacy Among Compliant Patients in a Randomized Clinical Trial with a Continuous Outcome

✱Afisi S. Ismaila, McMaster University, Dept of Clinical Epidemiology and Biostatistics, McMaster University, HSC-2C10, Hamilton, ON L8N 3Z5 Canada, *ismailas@mcmaster.ca*; Stephen Walter, McMaster University

Key Words: compliance, efficacy, randomized controlled trials, treatment preference

When there are departures from assigned treatments in a randomized controlled trial (RCT), researchers are faced with the problem of estimating treatment benefits in compliant patients while respecting randomization. Traditional intention-to-treat (ITT) analysis avoids biases associated with the alternative approaches like per protocol (PP) and as-treated (AT) analyses but it provides incomplete information about the expected treatment effect among patients who are committed to taking the treatment. In this talk, we examine four patterns of departures from assigned treatments and present the preference-based (PB) approach to estimating treatment benefits by extending the work of Walter et al. (2006) on binary data. Specifically, we develop new PB estimators for continuous outcomes for each pattern and address the statistical properties of these estimators.

Analysis of Skewed Distribution: Transformation or Robust Regression

✱Hongwei Wang, Merck Research Laboratories, Rahway, NJ 07065, *hongwei_wang@merck.com*; Arvind K. Shah, Merck Research Laboratories

Key Words: skewed distribution, log-transformation, rank-transformation, bootstrap, median, power

Encountering of positive data with skewed distribution is fairly common in many disciplines where the median is a preferred measure of central tendency and hence for inference. For comparison of medians, three strategies are considered: 1) Analysis of the rank-transformed data; 2) Analysis of the log-transformed data; 3) The least absolute error regression. The relative performance of these three methods is studied here with respect to their attained powers using bootstrap approach on a real dataset. The dataset comes from three similarly designed clinical trials (n=3083) with skewed distributed C-Reactive Protein as one of the secondary endpoints. For bootstrapping, various fractions of the data are repeatedly drawn with replacement on which the three methods are applied to test no difference among medians. The conclusions on utilization of the three methods are drawn and reported here.

Statistical Power Simulations on the Choice of Baselines in Clinical Trials

✱Din Chen, South Dakota State University, Department of Mathematics and Statistics, 1062 Western Ave, Brookings, SD 57006, *din.chen@sdstate.edu*; Pinggao Zhang, Global Biometrics

Key Words: Clinical trials, baseline, power analysis, ANCOVA

Multiple assessments of an efficacy variable are often conducted prior to the initiation of randomized treatment in clinical trials as baseline information. This paper is to investigate which baselines to be included in the analysis of covariance to increase the statistical power. A power analysis is developed with extensive simulations based on data from clinical trials in patients with end stage renal disease (ESRD). The findings can be easily applied in and extended to other clinical trials with similar design. Since a continuous efficacy variable is commonly dichotomized to a categorical variable for analysis and reporting in the medical society, the magnitude of the power loss is also examined.

Identification of Promising Subgroups in the Retrospective Analysis of Clinical Trials

✱Ilya Lipkovich, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *lipkovichia@lilly.com*; Alex Dmitrienko, Eli Lilly and Company; Eric Su, Eli Lilly and Company; Jonathan Denne, Eli Lilly and Company; Gregory Enas, Eli Lilly and Company

Key Words: Recursive partitioning, Classification trees, Data mining, Retrospective data analysis

Large amounts of data from clinical studies with a nonsignificant primary analysis are often underutilized and drug development programs are terminated early without a rigorous examination of subgroups of patients that could potentially benefit from the treatment. In this talk we propose a novel methodology for identifying such promising subgroups, based on a recursive tree partitioning algorithm. The proposed procedure incorporates a built-in mechanism for protection against false discovery by adjusting for multiplicity, tuning a complexity penalty for the data splitting criterion using cross-validation, and replication across validation datasets. In some cases these protections may not only generate hypotheses, but test and confirm hypotheses to drive regulatory decisionmaking. The results of a simulation study are presented and an application to real clinical trials is discussed.

Identifying Endpoints and Analysis Methodologies in the Assessment of Treatment Effects on Nocturia Episodes in Patients with Overactive Bladders

Xiaojiang Zhan, Merck & Co., Inc; ✱Liang Chen, Merck & Co., Inc, 126 E Lincoln Ave, RY34-A316, Rahway, NJ 07065, *liang_chen2@merck.com*; Weili He, Merck Research Laboratories

Key Words: overactive bladder, nocturia, recurrent events, GEE model, Longitudinal linear model

In the treatment of postmenopausal women with Overactive Bladder (OAB), a benefit to nocturia could be important for a novel OAB drug. However, clinical endpoints and statistical methodologies have not been systematically identified and investigated for the assessment of treatment effects on nocturia episodes. Several nocturia endpoints, such as averaged daily number of nocturia episodes/8hours and number of nocturia episodes per diary week, and a few statistical models, including the longitudinal linear model and GEE model, were proposed and investigated. Simulation studies were conducted to evaluate the properties of these endpoints and the statistical models. In the presentation, we will discuss the endpoint selection for nocturia symptoms and propose analysis approaches for these endpoints.

110 Nonlinear Models in Drug Development ●▲

Biopharmaceutical Section, Biometrics Section
Monday, August 4, 8:30 a.m.–10:20 a.m.

Clinical Statisticians' New Challenge: Pharmacokinetic/Pharmacodynamic Modeling and Simulation

✱Kuenhi Tsai, Merck & Co., Inc, UG1D-44, North Wales, PA 19454, *kuenhi_tsai@merck.com*; Alan Hartford, Merck & Co., Inc; Huafeng Zhou, Merck & Co., Inc.; Michael Hesney, Merck & Co., Inc

Key Words: Pharmacokinetic/Pharmacodynamic, Modeling and Simulation, clinical statistician

Pharmacokinetic/pharmacodynamic (PK/PD) modeling and simulation (M&S) has become important in early and late stages of drug development. In general, clinical statisticians are not involved in PK/PD M&S and have not been trained in the details of its work. For example, clinical statisticians analyze dose-response data, but do not extend the analysis to dose-exposure-response. Furthermore, clinical statisticians generally find themselves removed from the decisionmaking of dose selection and may produce empirical models which conflict with PK/PD models. To overcome this barrier, we will illustrate the general frame work of PK/PD M&S from preclinical to late-stage clinical development using case studies. We will also discuss what clinical statisticians may want to learn and new research areas in PK/PD M&S, such as physiologically based models.

On the Use and Challenge of Emax Model for Assessing Drug Interaction in Combination Studies

✱Diane D. Liu, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Unit 447, Houston, TX 77030-4009, *dianeliu@mdanderson.org*; Heather Y. Lin, The University of Texas M.D. Anderson Cancer Center; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center

Key Words: Drug combination, Emax model, drug interction, Ineraction Index, Loewe additivity

Drug combinations hold potential of increasing efficacy while reducing toxicity and have been successfully implemented in treating many diseases. Emax model can be applied to describe dose-response relationship. When cell growth is a measure of treatment effect, insufficient data can pose challenges in fitting the Emax model such as in situations where the measurement does not reach the maximum cell growth at low doses, no measurement is available in middle doses, or the measurement does not reach a plateau in high doses. In this paper, we give the conditions for Emax model to provide adequate model fit and offer remedies including standardizing and extending data to meet the model assumptions. Subsequently, calculations of interaction index and its variance are also given under the Loewe additivity. Simulations and case studies will be given to illustrate the proposed methods.

An Algorithm for Warfarin Dosing Using Nonlinear Mixed Effects Models

✱Kerrie Nelson, Massachusetts General Hospital, Biostatistics Unit, 50 Staniford Street Suite 560,, Boston, MA 02114, *kerrie@stat.sc.edu*; David Schoenfeld, Massachusetts General Hospital

Key Words: warfarin, doses, nonlinear mixed effects, algorithm

Warfarin is commonly prescribed for the prevention of blood clotting for many medical conditions. It has continually proved to be a challenge to determine doses for a patient due to wide intra- and inter-variability in its effects, where over- or underdosing can have serious consequences such as severe bleeding. Many maintenance dose algorithms have been developed to determine starting doses for new patients in order to achieve stability of the drug's effects in an individual in the shortest possible time. In this talk we propose a statistical model for describing the association between warfarin dosing and a patient's INR value, and present a simple algorithm incorpating genetic and other factors for determining a starting and maintenance doses to achieve stability for a patient. Results are based on a study being undertaken at Massachussetts General Hospital.

Using Genetic Algorithms To Find Starting Values for Fitting Nonlinear Models

✱Haiyuan Zhu, Merck Research Laboratories, 126 E Lincoln Av., RY34-A304, Rahway, NJ 07065, *haiyuan_zhu@merck.com*

Key Words: Nonlinear Models, Starting Values, Genetic Algorithm

In statistical modeling with nonlinear mixed effect models, specifying proper starting values for iterative model fitting algorithms is important. In practice, however, to set up those starting values could be difficult in some situations. Bad starting values could lead the fitting algorithms to a local maximum rather than the global one, or even fail to converge. When the fitting algorithm converges, it is not easy to verify whether it converged to global maximal point, or just trapped to a local maximal point. In this paper, a genetic algorithm (GA) will be used to find starting values for conventional model fitting algorithm. The values found by the GA could also serve as a probe to check whether the conventional algorithms converge to global maximum or not. An example of modeling the pharmacokinetics with a nonlinear mixed effect model will be used to illustrate the procedure.

Functional Data Analysis for ECG T-Wave Modeling

✱Yingchun Zhou, National Institute of Statistical Sciences, 19 T.W.Alexander Drive, Research Triangle Park, NC 27709, *zhouyc@niss.org*; Nell Sedransk, National Institute of Statistical Sciences

Key Words: functional data analysis, T wave, QT interval

The T-wave of an electrocardiogram (ECG) represents ventricular repolarization, and its changes reflect various cardiovascular conditions. Links between abnormal cardiac repolarization and malignant arrhythmias have been established. To study cardiac function through T-wave variation, functional data analysis is used to build a statistical model of T-wave shape. This four-parameter model explains approximately 95% of the variation in T-wave shape for normal subjects under ordinary conditions. It also distinguishes arrhythmias from normal heart rhythms on ECGs. Parameterizing the T-wave morphology is particularly useful for characterizing the repolarization process and any perturbations or changes over time. An important advantage of this approach over QT measurements is that it is not sensitive to the determination of the start and the end point of the T-wave.

A New Approach for Finding Global Minima in Nonlinear Least Squares Regression

✱Leonid A. Khinkis, Canisius College, 2001 Main Street, Department of Mathematics and Statistics, Buffalo, NY 14208, *khinkis@canisius.edu*; Milburn E. Crotzer, Canisius College

Key Words: Nonlinear Regression, Parameter Estimation, Confidence Intervals

Traditional methods for nonlinear regression analysis can lead to unusual estimates of the parameters of a model and their associated confidence limits. A method has been developed to determine whether a global minimum has actually been achieved in these cases, and to identify alternate parameter estimates associated with the unusal estimates. This technique improves the precision of the estimates. Applications of this method to models commonly used in the pharmaceutical and environmental sciences will be presented.

111 Model-Fitting ●

Section on Statistics in Epidemiology, Biometrics Section

Monday, August 4, 8:30 a.m.–10:20 a.m.

Finding Factors Influencing Risk: Comparing Variable Selection Methods Applied to Logistic Regression Models of Cases and Controls

✱Michael D. Swartz, The University of Texas M.D. Anderson Cancer Center, Department of Epidemiology Unit 1340, Houston,, TX 77030, *mdswartz@mdanderson.org*; Robert K. Yu, The University of Texas M.D. Anderson Cancer Center; Sanjay Shete, The University of Texas M.D. Anderson Cancer Center

Key Words: Bayesian logistic regression, Case-control analyses, Logistic regression, prior calibration, Variable selection

How one selects the factors to include in a disease risk model can impact results. This study compares the performance of several variable selection techniques applied to logistic regression. We performed realistic simulation studies to compare five methods of variable selection: a confidence interval approach for significant coefficients, backward selection, forward selection, stepwise selection, and Bayesian stochastic search variable selection using both informed and uniformed priors. We defined our simulated diseases

using covariate distributions, correlations and odd ratios for cancer risk found in the literature for environmental, nutritional and genetic risk factors and interactions. We used a permutation method to calibrate the priors of the Bayesian method and evaluate its sensitivity. We compared each method using average power and average false positive frequencies.

Bayesian Analysis of Covariate Profiles

✱John T. Molitor, Imperial College, London, Division of Epidemiology, Public Health and Primary Care, Lecturer in Statistics, London, W2 1PG United Kingdom, *john.molitor@imperial.ac.uk*; Michail Papathomas, Imperial College, London; Sylvia Richardson, Imperial College, London

Key Words: profiles, children's health, correlated data, bayesian analysis, mcmc

Standard regression analyses are often plagued with problems that occur when one tries to make meaningful inference using datasets that contain a large number of correlated variables. In this manuscript, we propose an inferential data analysis method that uses, as its basic unit of inference, a profile, formed from a sequence of covariate values. The model presented is based on Bayesian partition models. Our implementation of this approach extends the standard partition model in a number of important ways, such as, a) allowing number of clusters to be random, b) performing variable selection, and c) utilizing a set of post-processing procedures to provide an examination and comparison of different partitions of the data. An analysis of children's health data from The National Survey of Children's Health (NSCH) is provided.

Variable Selection in Linear Mixed Model: A New Algorithm Incorporating Investigator Preference and Nonmissingness of Data

✱Abu Minhajuddin, The University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Room E1.401, Dallas, TX 75390, *Abu.Minhajuddin@UTSouthwestern.edu*; Hrishikesh Chakraborty, RTI International

Key Words: Variable selection, statistical model building, linear mixed model, stepwise, forward selection, backward elimination

Variable selection in the context of a linear model or a linear mixed model is a fundamental but often a contentious part in the applied statistical model building. However, very little on the topic is available in statistical literature. In the current article, we propose a new algorithm for variable selection in the context of a linear mixed model that considers investigator preference and data availability along with other statistical consideration. The performance of the new algorithm is contrasted with the available automated variable selection approaches like stepwise, forward selection, and backward elimination and the best subset selection using a real data set. Cross-validation method is used to assess the predictive performance of the estimated model.

'Treatment - Outcome Complex' and Analysis of Observational Data

✱Lev S. Sverdlov, Schering-Plough Research Institute, 50 Lawrence Rd., Springfield, NJ 07081, *lev.sverdlov@spcorp.com*

Key Words: observational data analysis, Treatment-Outcome Complex, non-random assignment bias, cross-sectional analysis, longitudinal analysis, inference

The major obstacles towards making inferences from analysis of effects of drugs using observational data include a bias due to non-random assignment of patients to treatment modalities, which theoretically cannot be eliminated, and uncertain dependency relationships between analyzed variables. At any moment during treatment of chronic conditions, we iassess neither a severity of the disorder, nor an outcome of treatment. Rather it is a momentary functional state of the patient under the influence of therapy ("Treatment-Outcome Complex" - TOC). Analysis of cross-sectional and longitudinal patterns of TOC in the framework of Non-Random, RTC, Random, Titration, Incomplete Titration, and Mixed Models explores additional information to characterize relationships between treatment and outcome. The inferential opportunities, as well as limitations of this approach will be discussed.

Reducing Bias in Observational Studies Using Propensity Score Methods

✱Binod Neupane, McMaster University, Dept. of Clinical Epidemiology and Biostatistics, 1200 Main Street West, Hamilton, ON L8N 3Z5 Canada, *neupanbp@mcmaster.ca*; Lehana Thabane, McMaster University; Mark Loeb, McMaster University

Key Words: Observational studies, Bias, Data analysis, Propensity score methods

Multivariable regression adjustments and propensity score methods are two popular methods of reducing biases in the results of observational studies. Some theoretical and simulative studies suggest that the propensity score methods outperform multivariable regression adjustments in reducing biases under variety of situations. Also, some carefully conducted propensity score analyses confirm the results from randomized experiments. In this talk, we will present a review of arguments about how a propensity score model should be specified for unbiased analysis. We will present a brief review of different propensity score techniques and comparisons of propensity score methods with multivariable regression techniques in terms of reducing biases. Finally, we will present the result of empirical application of the methods to real data sets.

Length Optimal Interval Estimators for Multicenter Trials

✱Peng Zhang, Harvard University, 02115, *pzhang@hsph.harvard.edu*; Hajime Uno, Harvard University; Lee-Jen Wei, Harvard University

Key Words: optimal, confidence interval, multi-center, convex optimization

Abstract In epidemiology studies, we often face to analyze multicenter trials data. For the trial with low variability, the interval estimator changes steadily with the nominal coverage level. On the contrary, for the trial with high variability, the interval estimator can change dramatically with the nominal coverage level. In this paper, we propose to use the total length of all confidence intervals as a global assessment of the sequence of interval estimators, while maintaining the marginal coverage level. By convex optimization theory, we first develop an efficient algorithm to obtain the minimizing coverage level for each trial, and then prove that it will give the correct coverage. We demonstrate our method through data sets from a multicenter clinical study and also from a typical microarray experiment. The performance of our procedure is examined via simulation studies.

Stochastic Dynamics and Dynamical Systems in Biosystem Modeling

✱Huitian Lu, South Dakota State University, SD 57007, *huitian.lu@sdstate.edu*; Jin Kai, Texas A&M University, Kingsville; Amy Wang, Golden Computing Technology Solutions

Key Words: dynamical systems, stochastic dynamics, mathematical biology, cancer modeling

The paper discusses the modeling of nonlinear dynamical behavior of biosystems in both deterministic and stochastic considerations. How dynamical systems and both dynamical systems and stochastic analysis work on the modeling. Bifurcation and equilibrium stability are also discussed. The stochastic dynamical features of cancer stem cells are studied and explored in a system view as well.

112 Invited Poster Presentations: Class Projects

Section on Statistical Education
Monday, August 4, 8:30 a.m.–10:20 a.m.

Service Learning Project: Connecting Student Needs with the Community

Neal Rogness, Grand Valley State University; ✱Paul Stephenson, Grand Valley State University, Department of Statistics, 1133 MAK, 1 Campus Drive, Allendale, MI 49401, *stephenp@gvsu.edu*

Key Words: service-learning, community, project

In the summer of 2007, the Supervisor of Allendale Township (the township in which Grand Valley State University is located) initiated a conversation with GVSU's Statistical Consulting Center (SCC) regarding a survey of students. The objective of the survey was to obtain information from GVSU students on services offered by Allendale Township and businesses. The authors, the elected officials in Allendale, several GVSU Student Senate leaders, and the SCC student consultant collaboratively developed a survey that was taken by students in select general education courses. The results of the survey were analyzed by students in STA 319 (Statistics Project) and were presented to the Allendale Township Board both in written and oral formats. The authors will present an overview of this initiative and will be prepared to discuss some of the benefits of this service learning experience.

Class Project: Service-Learning in Undergraduate Survey Sampling

✱John Gabrosek, Grand Valley State University, 1 Campus Drive, Dept of Statistics, Allendale, MI 49401-9403, *gabrosej@gvsu.edu*

Key Words: service-learning, survey sampling

The author has incorporated a service-learning component into an introduction to survey sampling course designed for statistics majors and applied statistics minors. Issues related to introducing service-learning into a statistics course are investigated. Discussion of the ups and downs is presented. Details of the service-learning results are briefly mentioned.

Class Project: Measuring Student Perceptions of the General Education Program

✱Whitney B. Miner, Grand Valley State University, Department of Statistics, 1 Campus Drive, Allendale, MI 49401, *whit04@comcast.net*; Phyllis J. Curtiss, Grand Valley State University

Key Words: Class Project, assessment, general education

Introductory statistics students designed questionnaires to determine how Grand Valley State University students felt about the general education program. Each group decided on their own questions, with some guidance from the professor. After collecting the data, a short analysis was provided. The data sets were analyzed by a nonparametric class the following semester. Results are being tabulated and will be presented to the General Education Assessment Committee at GVSU.

113 Contributed Poster Presentations

Section on Bayesian Statistical Science, Section on Physical and Engineering Sciences, Section on Statistical Consulting, Section on Statistical Education, Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences
Monday, August 4, 8:30 a.m.–10:20 a.m.

Bayesian Modeling of Wind Fields Using Surface Data Collected Over Land

✱Margaret Short, University of Alaska Fairbanks, P.O. Box 750125, Fairbanks, AK 99775, *jsm@mshort.authorized.yon.net*

Key Words: wind field, process convolution, Markov chain Monte Carlo

We propose an approach to modeling wind field in which the error structure includes the type of instrumentation used to collect such data over land, namely anemometers (for wind speed) and vanes (for wind direction). Thus the model can handle both the periodicity of the wind direction and the nonnegativity of the wind speed. The measurement error depends in part on the wind speed; this characteristic is incorporated in the model. We use a Bayesian process convolution approach, implemented via Markov chain Monte Carlo. We illustrate the model performance with surface wind data set from Fairbanks, Alaska. We present initial results incorporating topography in our model.

Impact of a Regional Trend on the Estimation of Unconventional Natural Gas Resources Using Local Nonparametric Prediction Models

✱Timothy C. Coburn, Abilene Christian University, ACU Box 29315, Abilene, TX 79699, *tim.coburn@coba.acu.edu*; Emil D. Attanasi, U.S. Geological Survey; Phillip A. Freeman, U.S. Geological Survey

Key Words: Local spatial nonparametric prediction models, Unconventional natural gas resources, Cross-validation, Jackknife, Boostrap

This presentation extends the results of an earlier investigation of the use of local nonparametric prediction models for estimating the remaining volumes of hydrocarbon resources at untested locations, with particular reference to continuous-type natural gas plays. Previous work determined that an approach that marries cross-validation, jackknife, and bootstrap procedures yields reliable estimates of regional totals as well as the amounts at localized individual sites. However, no underlying spatial trend was assumed. The present study incorporates an adjustment proposed by Tomczak in 1998 to account for the presence of a regional spatial trend in the subsurface. Simulation indicates there is, indeed, an overall improvement in the mean square errors associated with competing methods, but that the ranking of these approaches in terms of predictive performance is not altered.

Geographic Variability of Asthma Prevalence in Inner City Schools: Chicago Initiative To Raise Asthma Health Equity

✱Laura M. Curtis, Northwestern University, 750 N Lake Shore Drive, 10th Floor, Chicago, IL 60611, *l-curtis@northwestern.edu*; Xingyou Zhang, The Robert Grahm Center; Ruchi S. Gupta, Northwestern University & Children's Memorial Hospital; Kevin B. Weiss, Northwestern University

Key Words: asthma prevalence, geographic variability, elementary schools, children

Asthma is one of the most common chronic illnesses in public and private inner city schools. Multilevel regression models were used to evaluate what factors are associated with higher prevalence of asthma in 105 schools in Chicago, including age, gender, and race/ethnicity at the individual level and school environment measured by demographic and socioeconomic composition, year built, and type of school (public vs. Archdiocese). This study seeks to characterize the geographic variability of asthma prevalence among these schools and determine whether any of the above characteristics account for this variability. Different spatial cross random effect mixed models, including spatial random effects for spatial variation and the school random effects for school variation will be used and compared.

Effective Collaboration: Are Junior Faculty Knowledgeable About How To Work with Statisticians?

✱Brianna C. Bright, University of Oklahoma Health Sciences Center, 801 NE 13th St., Room 309BA, Oklahoma City, OK 73104, *Brianna-Bright@ouhsc.edu*

Key Words: collaboration, consulting, survey, junior faculty

Multidisciplinary research teams are common in medical centers and often require a statistician. A 12-question survey, administered to 19 junior faculty members across many disciplines in a health sciences center, assessed perceived grasp of statistical concepts and practical collaboration knowledge. The survey showed that more education is necessary to facilitate effective collaboration between junior faculty and statisticians. Multiple gaps in the practical collaboration with statisticians exist: 37% didn't know how to get in contact with a statistician; 56% believed that statisticians are readily available to provide free services. 95% knew to contact a statistician before collecting data, though almost all were working on a research project and only 33% had contacted a statistician. 32% cited their lack of knowledge of how to work with a statistician as a barrier to getting help.

Tips for Providing Statistical Consultation to Medical Trainees

✱Christina A. Haines, Biostatistics Consulting Group, Department of Pediatrics, University of Alberta, Rm 9425, Aberhart Centre, 11402 University Ave NW, Edmonton, AB T6G 2J3 Canada, *haines@ualberta.ca*; Rhonda Rosychuk, University of Alberta

Key Words: Biostatistics, Consulting, Medicine, Trainee, Strategies, Success

Beginning a career as a biostatistics consultant in academic medicine requires adaptation and patience as many new situations are faced. Consulting medical trainees (i.e. residents, fellows) with a wide range of research experience and attitudes towards research further presents unique challenges and opportunities. Often, trainees have not previously consulted a biostatistician and, with varied supervisor-trainee dynamics, must complete a research project conforming to program expectations, usually with limited timelines and budgets. Establishing clear communications, instructing on good data procedures, specifying timelines and tasks, and encouraging active participation in analyses are some techniques leading to positive trainee-statistician relationships. This poster will further describe the approaches used by a new consultant to ensure successful trainee research experiences.

Projects STATCOM UHasselt

Amparo Y. Castro, Hasselt University; ✱Herbert Thijs, Hasselt University, Agoralaan - Building D, Office D5, Diepenbeek, International 3590 Belgium, *herbert.thijs@luc.ac.be*; Daniel Martinez, Hasselt University; Vishva Danthurebandara, Hasselt University

Key Words: STATCOM network, statistical consulting, pro bono statistics, UHasselt

The objective is to present two of the projects in which Statcom Uhasselt has been involved. In the first one, the aim is to establish the association between the hyaluronic acid binding assay (HBA) and the success of an IVF and ICSI treatment. Correlation analysis, multiple linear regression, logistic regression and association tests were used to achieve the project goals. The second project is "Induction of oxidative stress related responses in Arabidopsis thaliana following uranium exposure." The objective is to find differences between a control group and four uranium dose levels, in the absorption of uranium and the presence of several enzymes, minerals, metabolites, leaves growth, root growth and DNA concentration. The statistical analysis involves non parametric Kruskall Wallis, ANOVA, random coefficient models, continuous marginal models, and nonlinear surface response models.

STATCOM at Chicago: Universities Working Together To Serve the Community

✱Meredith Wascher, DePaul University, Interdisciplinary Science & Technology Center, 990 W Fullerton Ave, Suite 4400, Chicago, IL 60614, *mwascher@depaul.edu*; Bart Phillips, E-B Research, LLC; Diya Zhang, Northwestern University

Key Words: STATCOM network, statistical consulting, pro bono statistics, STATCOM

The Chicago Chapter of the ASA is bringing together the Chicago area's many universities and members of the professional statistical community to form a collaborative STATCOM at Chicago. This effort is the first of its kind in the STATCOM Network. The group is working to set up an infrastructure that supports two ultimate outcomes: first, to become a visible and established service organization in Chicago; and second, to facilitate communication and build a community among students at neighboring universities. This presentation discusses the framework for a multi-university STATCOM, the benefits and challenges of the collaborative model, and ideas for the future.

Seven Years of STATCOM at Purdue University: Managing a Growing Number of Student Volunteers

Douglas Baumann, STATCOM at Purdue; ✱Andrea Rau, Purdue University, Department of Statistics, 250 N. University Street, West Lafayette, IN 47907, *STATCOM@stat.purdue.edu*

Key Words: STATCOM network, statistical consulting, pro bono statistics

Statistics in the Community (STATCOM) at Purdue is a volunteer community outreach organization directed and staffed by graduate students which provides free professional statistical consulting services to governmental and nonprofit groups. Since its establishment in 2001, the STATCOM student volunteer base at Purdue has grown to over 55 students across several departments. In addition to growing its volunteer base, STATCOM at Purdue has developed new ways to serve the community, such as P-12 Outreach and STATCOM Network components. This growth in student participation and available services has required STATCOM at Purdue to develop a broader organizational infrastructure, adding the roles of Project Manager, Network Outreach Coordinator, and P-12 Officer. This enhanced structure allows students to develop strong leadership skills by serving in a variety of different organizational roles.

STATCOM at Iowa State: Experiences and Projects in the First Year

✱David Rockoff, Iowa State University, 240 Raphael Ave. #19, Ames, IA 50014, *drockoff@iastate.edu*; Jennifer C. Huckett, Iowa State University; Anna Peterson, Iowa State University; Jessica Chapman, Iowa State University; Nicholas Beyler, Iowa State University; Jonathan Hobbs, Iowa State University; Timothy Bancroft, Iowa State University

Key Words: STATCOM network, statistical consulting, pro bono statistics

Statistics in the Community (STATCOM) at Iowa State joined the STATCOM Network in 2007. Graduate students involved in STATCOM volunteer their time to provide pro bono statistical consulting to local nonprofit and government organizations in the community. Consultants work in teams to enable clients to address specific questions, enriching their organization and the community. In addition to gaining experience in statistical consulting, participation in STATCOM develops members' skills in leadership, management, and communication. Our current clients include the Volunteer Center of Story County and the Retired Seniors Volunteer Program. We will present details on two consulting projects and will be available to share our experiences and challenges in launching STATCOM at Iowa State University.

STATCOM at the University of Washington: Implementing a Peer Evaluation System for a Local Fire Department

✱David W. Lockhart, University of Washington, 325 Ninth Ave, Box 359931, Seattle, WA 98104, *dwl@u.washington.edu*; Julian Wolfson, University of Washington; Annette Ghee, University of Washington; Rori Rohlfs, University of Washington

Key Words: STATCOM network, statistical consulting, pro bono statistics, survey methods, missing data, mixed effects model

Statistics in the Community (STATCOM) at the University of Washington (UW) was founded in 2005. One of the largest and longest-running projects we have participated in so far involves a local fire department serving a community of approximately 50,000 north of Seattle. After some initial experience developing a peer evaluation system, fire department representatives approached UW STATCOM with questions about how to best analyze these data and assess potential biasing effects. Ultimately, we became a partner in refining the entire peer evaluation procedure. We have provided support on questionnaire design including acceptability of the changes to the system, sampling to select raters, and data analysis. In this poster, we provide an overview of UW STATCOM activities and describe some of the more interesting features of the fire department project.

STATCOM at University of Michigan: Making a Michigan Difference - A Promising Beginning

✱Maria Larkina, The University of Michigan, 1420 Washington Heights, Biostatistics, Ann Arbor, MI 48109-2029, *mlark@umich.edu*

Key Words: STATCOM network, statistical consulting, pro bono statistics, University of Michigan

STATCOM at the University of Michigan student group has been officially recognized for just one year, here is a short overview of the hurdles a new student group had to overcome in order to deliver free statistical services to the community. From student recruitment to advertisement in the community, the effort to convince local organizations of how statistics can be useful to them. Our first several projects have been slow to develop, and have taught us valuable lessons in statistician/client interactions, the challenge of communication with non-profits and government offices. We also had to meet the demand of translating between statistical jargon and lay terms in order to formulate projects into manageable problems for the students. This poster is intended to share our enriching experience with other young student groups providing statistical consulting to general population.

Classroom Use of R: Coverage Probabilities of Poisson Interval Estimates

✱Lioudmila Belan, California State University, East Bay, 707 San Conrado Terr #7, Sunnyvale, CA 94085, *mila.belan@gmail.com*; Bruce Trumbo, California State University, East Bay; Eric Suess, California State University, East Bay

Key Words: Confidence interval, Bayesian probability interval, Poisson mean, R/S-Plus, Pedagogy

True coverage probabilities of several nominal 95% and 99% interval estimates for the Poisson mean are computed using elementary programs in R. Among the intervals considered are frequentist confidence intervals based on normal approximation and Bayesian posterior probability intervals resulting from a noninformative prior.

Real Data with Real Questions: Computer Labs for Introductory and Intermediate Statistics

✱Rosemary Roberts, Bowdoin College, Department of Mathematics, 8600 College Station, Brunswick, ME 04011-8486, *rroberts@bowdoin.edu*; Ann Cannon, Cornell College

Key Words: computer labs, classroom activities, introductory statistics

We have developed a suite of five computer labs to enhance the teaching of both the introductory statistics course and an intermediate statistics course. Each of the labs is implemented for two or three moderate size data sets—one on Savannah Sparrows, one on various species of Hawks and one investigating the use of the pesticide fenthion on olive groves in Greece. Students are led, step by step, through the analysis using DataDesk or Minitab. In addition, each lab also has both an exploratory and a writing extension that the students can be asked to complete outside of class. The labs, extensions, model solutions, and data sets are freely available on the web.

Teaching Introductory Statistics with Simulations in JMP Statistical Discovery Software

✱William M. Duckworth, Creighton University, 2500 California Plaza, Omaha, NE 68178, *wmd@creighton.edu*; Amy Froelich, Iowa State University; Mark Bailey, SAS Institute Inc.; Wayne Levin, Predictum Management Sciences, Inc.

Key Words: classroom activities, applets, scripting, scripts

Simulation has become an important tool in teaching topics related to sampling distributions and inference in the introductory statistics class. Many of these simulations have been developed with Java applets and made available on the web. While these applets are easy to use and readily available to statistics instructors, they may not match classroom, laboratory or homework activities. As a result, students can struggle with the transition between the classroom and the computer simulation activities. We will showcase JMP scripts developed to take advantage of JMP's powerful interface features and to provide flexibility to match the simulation experience in JMP to classroom, laboratory or homework activities traditionally used in introductory statistics courses.

Teaching Undergraduate Statistics as a Capstone for Mathematics Majors

✱Jeremy Nadolski, Benedictine University, 5700 College Rd, 126 Birck Hall, Lisle, IL 60532, *jnadolski@ben.edu*

Key Words: education, capstone, undergraduate, statistics

In a classroom consisting of future high-school teachers, actuaries, mathematics graduate students, and other math-related fields, the course content must be a utility and resource for the varied student population. An

undergraduate mathematics major at BU must take an upper level sequence which involves one research project to serve as a capstone experience. During the second semester Probability and Statistics course, students focus on applying probability and statistical methods learned in the course to analyze data. The analysis of data is covered in exams and the students' research project. Students are asked to provide analysis on data sets, formally write up the analysis and orally present findings externally to peers and faculty. This poster will discuss the purpose and how to allow the students to gain hands-on experience as well as the benefit to the classroom.

A First Step to Understanding the Difficulty in Teaching Sampling Distributions, Sampling Error, and Statistical Inference

✱S. David Kriska, The Ohio State University, 3758 Surrey Hill Place, Fisher College of Business, Upper Arlington, OH 43220, *davidkriska@sbcglobal.net*; Mark C. Fulcomer, Richard Stockton College of New Jersey; Marcia M. Sass, University of Medicine and Dentistry of New Jersey (UMDNJ)

Key Words: Elementary statistics, sampling distributions, sampling error

Delmas, Garfield, Ooms, and Chance (2007) note that even after having taken an introductory statistics class, many students continue to struggle with the concepts of sampling distributions, sampling error, and statistical inference. As a first step in identifying ways of more effectively approaching these topics, the poster will examine methods currently used in popular text books and propose an alternative based on using binomial and Poisson distributions. The examination of textbooks will identify variability or lack of variability in the approaches to teaching these topics with particular attention being paid to the use of statistical language. The poster will also illustrate how software developed by Fulcomer (2007) can be used to formulate interesting exercises useful in teaching the three topics of sampling distributions, sampling error, and statistical inference.

Teaching Statistical Consulting

✱Lynn Eudey, California State University, East Bay, Department of Statistics and Biostatistics, 25800 Carlos Bee Boulevard, Hayward, CA 94542, *lynn.eudey@csueastbay.edu*

Key Words: consulting, pedagogy

In most masters' programs students are taught experimental design, the mathematical foundations and applications of commonly used statistical methods. Yet, the majority of professional positions requires effective consulting practices. The new statistician is expected to communicate technical concepts to audiences with varying statistical backgrounds. Many programs, especially doctoral programs have a course in statistical consulting. This poster will present an overview of the pedagogy used for training statistical consulting in universities and colleges in the United States.

A Hunt for Better Features in Teaching an Introductory Statistics Class

✱Julia A. Norton, California State University, East Bay, 28022 El Portal Drive, Hayward, CA 94542-2512, *julia.norton@csueastbay.edu*; Yan Yan Zhou, California State University, East Bay; Farnaz Ganjeizadeh, California State University, East Bay

Key Words: teaching, assessment, pedagogy

After many years of instruction and program assessment, it seemed time to concentrate on individual courses. Teaming two faculty members in Statistics and Engineering, we consider modifications including changes to reduce incorrect items in student responses on quantitative tests, increased writing to emphasize critical analysis, and increased personalized feedback. For statistics measurements compare testing early in the quarter versus later, item analysis of missed questions and problems, study of mistakes on projects and written assignments and relationships among these types of questions. In engineering the evaluation process of problem solving requires distinguishing "missing conceptions" evaluation of understanding the theories and "misconceptions" (making mistakes when applying the concepts). Effective evaluation can provide feedback based on the students' learning needs.

Your Next Teaching Vehicle May Be a Hybrid: Getting More Mileage from the MBA Quant Course

✱Patrick S. Noonan, Emory University, 1300 Clifton Road NE, Goizueta Business School, Atlanta, GA 30322, *patrick_noonan@bus.emory.edu*

Key Words: teaching, MBA, decision analysis, business statistics, statistical education

Among the many challenges faced by educators in our field is the shrinking real estate in MBA programs. Never has it been more important for organizations to have leaders who can think clearly about risk and uncertainty, and engage critically with vast amounts of available data. Likewise, never before have nontechnical managers had such access to analytical tools for exploring and modeling complex problems. However, at the same time we are presented with this enormous opportunity to both shape the mental models and stock the analytical toolkits of MBAs, we are being pressed, quite severely in some cases, to reduce our footprints in MBA curricula. Some programs have addressed this compression by creating a single-semester "quant" course. One approach to this is to create a "hybrid," combining stats and OR/MS. But the key to success is centering the course around decision analysis.

Using the GAISE Guidelines in a Graduate Biomedical Statistical Methods Class

✱Emily H. Sheldon, Virginia Commonwealth University, Biostatistics 980032, Richmond, VA 23298-0032, *SheldonEH@vcu.edu*; Jessica M. Ketchum, Virginia Commonwealth University; Al M. Best, Virginia Commonwealth University

Key Words: Statistical Methods, Teaching, Writing Intensive, Active Learning

The GAISE guidelines are recommended for teaching statistical literacy within a college setting. The poster will describe how we have addressed each of the GAISE guidelines in our graduate biostatistical methods class for biomedical researchers. In particular the course develops statistical thinking and literacy through active learning strategies such as problem-based and team-based learning. Homework and exams utilize real data and stress conceptual understanding through in-depth writing exercises. JMP software is used to provide students with a tool for calculation and allows focus to remain on comprehension rather than tedious formulas. Students are given feedback on their writing and opportunities to improve. In conclusion, as the GAISE guidelines have been increasingly incorporated into the coursework over the years students have demonstrated improved levels of learning.

A Visual Model for the Variance and Standard Deviation

✱James B. Orris, Butler University, College of Business Administration, 4600 Sunset Blvd., Indianapolis, IN 46208, *orris@butler.edu*

Key Words: standard deviaiton, variance, visual model, descrptive statistics

All introductory statistics courses discuss the variance and standard deviation; however, in many cases students just calculate a number using a computational equation but have little appreciation for what it represents. This approach focuses on the definitional equation for the variance that is based on the sum of squared deviations from the mean. It looks at each squared deviation as a graphical object, i.e., a square. A series of displays show how the standard deviation is the size of the 'average' square. This model gives students a better appreciation for the variance and standard deviation since they can visualize it graphically. This model may also be expanded in the future to represent more advanced concepts, perhaps a graphical analysis of variance.

A Probabilistic Look at a Calculus Problem

✱Lewis VanBrackle, Kennesaw State University, 1980 Towne Manor Drive, Kennesaw, GA 30144, *lvanbrac@kennesaw.edu*

Key Words: Calculus, Random Variable, Simulation

The problem of finding the minimum length of the crease formed when folding a sheet of paper so that the bottom left corner touches the right edge of the sheet appears in many calculus texts. When the sheet is folded so that the bottom right corner touches the left edge at a random point, finding the distribution of the length of the crease is an interesting probability problem. Solving this problem is a good exercise for students, requiring them to use inverse functions, functions of a random variable and simulation. This poster presents the steps involved in solving the problem and the results for 8.5 inch by 11 inch and 8.5 inch by 14 inch sheets of paper.

Two-Sample Considerations: From AP Question to Student/Faculty Research

✱Ken Constantine, Taylor University, Department of Mathematics, Upland, IN 46989, *knconstantine@tayloru.edu*

Key Words: two-sample, AP Statistics, post-stratification

In this poster, we explore an Advanced Placement Statistics Problem (#4, 2006 free response exam) which is a classic two-sample comparison of means. We briefly present the published solution and then critique the assumptions on which it is built, especially the assumption that independent random samples were taken. We use standard probability tools and simulation work to address these questions about the independence assumption: (i) Were the samples really independent? (ii) How does this assumption affect the methods used to obtain the published solution? (iii) Were the sample means uncorrelated? (iv) Are the results from the classical methods reliable? This example offers multiple classroom activities for student exploration of: sampling distributions, distinction between random variables which are independent and those which are uncorrelated, and simulation work.

Random Structures: A Course Bridging Statistics and Mathematics Programs in the Liberal Arts

✱Brian D. Jones, Kenyon College, 303 Hayes Hall, Gambier, OH 43022, *jonesbd@kenyon.edu*

Key Words: probability, education, random graph

The author has written a new course that serves both the statistics and mathematics programs at Kenyon College. This course explores the theory, applications, and interesting consequences when probability is introduced upon various mathematical objects. Some of the core topics are random graphs, random walks, and random permutations, as well as randomness applied to polynomials, functions, integer partitions, groups, and codes. The focus of the course is solving challenging problems and working on small projects. In addition to studying the random structures themselves, a concurrent focus of the course is the development of mathematical tools to analyze them, such as combinatorics, indicator variables, generating functions, laws of large numbers, asymptotic theory, elementary linear algebra, and computer simulation. The prerequisite for the course is only second-semester calculus.

Active Learning: Bayesian Statistics Modules You Can Include in Your Frequentist First or Second Statistics Course

✱Linda B. Collins, The University of Chicago, Department of Statistics, 5734 S. University Ave., Chicago, IL 60637, *collins@galton.uchicago.edu*

Key Words: Bayesian statistics, active learning, introductory statistics, projects

Almost all colleges now offer a “Stat 101” course. Some even offer a “Stat 201” data-based introduction for mathematically inclined students. In either course, the material tends to be essentially frequentist: calculating p-values and confidence intervals. We present active-learning-style modules introducing Bayesian statistics useful in any 2nd course with “Stat 101” as prerequisite or as a final project in a “Stat 201” course. We begin with a guided walk through a motivating Bayesian example, making comparisons to frequentist methods students would already be familiar with (e.g., confidence vs. Bayesian probability statements). After collecting and organizing data and using Bayesian techniques for the analysis in the first module, a second module has more examples, but is less prescribed and can take the form of a project students could complete largely on their own.

Classroom Illustrations of the Construction of Discrete Sampling Distributions

✱Mark C. Fulcomer, Richard Stockton College of New Jersey, 48 Trainor Circle, Bordentown, NJ 08505, *mcfulcomer@aol.com*; S. David Kriska, The Ohio State University; Marcia M. Sass, University of Medicine and Dentistry of New Jersey (UMDNJ) ; Maritza Jauregui, Richard Stockton College of New Jersey

Key Words: teaching statistical concepts, sampling distributions, computational simplifications

Despite being an important concept in statistical theory, sampling distributions are difficult to convey in introductory courses. In part these difficulties stem from preliminary topics (e.g., probability theory and statistical independence) that must first be mastered. However, even with simple discrete random variables, the rapid expansion of the calculations needed to visualize results as sample sizes increase can quickly introduce not only obstacles for student learning but also teaching challenges for instructors. Using readily available software to overcome computational issues, this presentation describes the construction of sampling distributions for two “competing” discrete processes in forms suitable for a single lecture to illustrate key concepts (e.g., unbiasedness). An example demonstrates an application involving small sample sizes (e.g., disease outbreak investigations).

Incorporating Medical Literature in the Statistics Classroom: From Initial Idea to Execution

✱Kirk Anderson, Grand Valley State University, 1 Campus Drive, Allendale, MI 49401, *anderkir@gvsu.edu*

Key Words: medical literature, health sciences, statistics education

If you are considering the use of medical literature in your classes, view this poster with examples from my experience. Various approaches will be explored from the initial thought, finding an appropriate paper, and making use of it in class. I am a statistics professor with experience teaching introductory statistics to graduate students who are typically biologists, nurses, occupational therapists, physical therapists, or physician assistants.

A Synergistic Effort To Recruit Future Biostatisticians in Arkansas High Schools

✱Songthip Ounpraseuth, University fo Arkansas for Medical Sciences, 4301 W. Markham St. Slot 781, Little Rock, AR 72205, *stounpraseuth@uams.edu*

Key Words: AP Statistics, High School, Education

The objectives of this project were to raise awareness of the discipline of biostatistics and career opportunities in biostatistics to central Arkansas high-school students involved in the Advanced Placement (AP) Statistics program. This was accomplished primarily by visiting several chosen central Arkansas high schools and giving creative presentations about biostatistics to high school students involved in the AP program. Additionally, the presentation slides and materials were composed into a teaching packet and distributed by mail to all remaining Arkansas public schools offering AP Statistics courses.

Recruiting Undergraduate Statistics Majors and Minors

✱Phyllis J. Curtiss, Grand Valley State University, Department of Statistics, 1 Campus Drive, Allendale, MI 49401, *curtissp@gvsu.edu*; Paul Stephenson, Grand Valley State University

Key Words: Recruitment, Statistics Majors, Curriculum

This poster session will give some ideas for recruiting undergraduate statistics majors and minors. We will discuss curriculum changes; as well as ideas and activities we have used at Grand Valley State University in order to recruit new majors and minors. GVSU is a division 2 school with a total enrollment of 23, 464 students. The university has seen a 33% increase in enrollment over the past ten years. Over this ten year period, the Statistics Department has increased the number of majors from 16 to 60 and the number of minors from 19 to 62. We currently have 14 tenure track faculty members in our department. We offer approximately 110 sections of introductory applied statistics per year, with a cap of 31 students per section.

Analysis of Sampled Fishbone Diagrams for Completeness and Size of a Second Stage

✱Frank Matejcik, South Dakota School of Mines & Technology, 501 E Saint Joseph St, Rapid City, SD 57701-3901, *frank.matejcik@sdsmt.edu*

Key Words: Fishbone, Ishikawa, capture, recapture

Procedures for creating a Fishbone diagram often assume all the possible participants contribute. However, sampling among the participants may be required. The sampled Fishbone diagrams are analyzed to estimate the portion of the Fishbone diagram that is completed in the sample. Additionally, recommendations for the size of a second sample to achieve a desired level of completeness are computed. The estimates use existing software for capture-recapture studies. Data was obtained in classroom activities. Additionally, a follow up activity is described. Possible applications in education, business, and government are presented, also.

114 *JASA* Theory and Methods Invited Paper Session

JASA, Theory and Methods, Section on Nonparametric Statistics, WNAR
Monday, August 4, 10:30 a.m.–12:20 p.m.

The Nested Dirichlet Process

✱Alan Gelfand, Duke University, Institute of Statistics & Decision Scien, Durham, NC 27708-0251, *alan@stat.duke.edu*; Abel Rodriguez, University of California, Santa Clara; David Dunson, National Institute of Environmental Health Science

Key Words: clustering, hierarchical model, nonparametric Bayes

In multicenter studies, subjects in different centers may have different outcome distributions. We discuss nonparametric modeling of these distributions, borrowing information across centers while also allowing centers to be clustered. Starting with a stickbreaking representation of the Dirichlet process (DP), we replace the random atoms with random probability measures drawn from a DP. This results in a nested Dirichlet process (NDP) prior, which can be placed on the collection of distributions for the different centers, with centers drawn from the same DP component automatically clustered. Theoretical properties are discussed. Additionally, an efficient MCMC algorithm is developed for computation. The methods is illustrated using a simulation study and an application to quality of care in U.S. hospitals.

115 Warranty and Other Field Reliability Data ●

Section on Physical and Engineering Sciences, Section on Quality and Productivity, SSC
Monday, August 4, 10:30 a.m.–12:20 p.m.

Field Reliability Improvement Through Analysis and Modeling of Sensor-Based Data: Opportunities and Challenges

✱Necip Doganaksoy, General Electric, GE GRC, 1 Research Circle, K15A54A, Niskayuna, NY 12309, *doganaksoy@research.ge.com*

Key Words: Automated monitoring, Data-mining, Multivariate modeling, Product reliability

The advances in sensor technologies have enabled manufacturers to track their products remotely and gather useful information to assess their health. The resulting data provides significant opportunities, as well as challenges, to statisticians in automated monitoring to identify impending failures. This is in sharp contrast to the recent past where field reliability data mainly comprised of after the fact failure information. Owing to large size and high dimension of sensor-based data, data-mining and multivariate statistical methods are particularly well suited to tackle the technical challenges. This presentation will focus on a recent application involving remote monitoring of locomotives for early detection of impending failures.

Using Experts' Knowledge To Forecast Warranty Claims

✱Marc Fredette, HEC Montreal, 3000 Côte-Sainte-Catherine, Montreal, QC h3t 2a7 Canada, *marc.fredette@hec.ca*

Key Words: Warranty Claims, Prediction, Random Effects Models, Poisson Processes, Prior Information

We discuss methods for predicting the eventual total number of warranty claims in a large population of units manufactured and sold over a period of time. As warranty claims data accumulate, methods based on Poisson processes with random effects often give appropriate frequentist prediction intervals. However, these methods fail to provide adequate intervals when they are based on early or limited data. Models' extensions incorporating experts' knowledge and other types of prior information will be discussed, and an application to the prediction of automobile warranty claims will be considered.

Warranty Analysis for Cost Reduction and Customer Satisfaction

✱Jeffrey A. Robinson, General Motors R&D Center, Mail Code 480-106-256, Warren, MI 48090-9055, *jeffrey.a.robinson@gm.com*

Key Words: problem solving, repair effectiveness, bill of material

Warranties are agreements between customer and manufacturer, but they are also product attributes to which customers assign value. Manufacturers naturally want to keep their costs low and customer perceptions of value high. Good basic engineering is obviously important in delivering first time quality. But for items that do come in for service under warranty, good diagnostics and highly effective repairs minimize the number of comebacks. Some automotive examples are presented in this talk to illustrate these issues. Some of the analysis methods include basic plotting and analysis of warranty claims data, examining variation in search of best practices, and primitive data mining to reveal patterns of repairs that might be useful in improving diagnosis. In some cases the methods supporting the examples are in routine use, while others are experimental.

Using Life Data To Assess the Risk of Product Failure

✱William Q. Meeker, Iowa State University, 304C Snedecor Hall, Ames, IA 50011, *wqmeeker@iastate.edu*

Key Words: Censored data, Competing risk, Prediction

Unanticipated failure modes are the primary cause of serious product reliability failures. After such a failure mode is discovered for a product that is already in the field, an immediate question that management needs to answer is “what is the financial risk of failures of product already in the field?” Risks can involve costs ranging to excessive warranty returns to loss of human life. Limited failure data, often with complications, perhaps supplemented by engineering information, are generally used to make the needed assessment. In this talk, I will discuss some well-known and some not-so-well-known examples of unanticipated failure modes and the risks involved. I will then outline some details of the statistical methods of life data analysis that were used to assess risk for one particular product.

116 Health/Biological Imprint of Climate and Atmospherics ●

Section on Statistics and the Environment
Monday, August 4, 10:30 a.m.–12:20 p.m.

Spatio-Temporal Threshold Models for Relating UV Exposures and Skin Cancer in the Central United States

Laura A. Hatfield, The University of Minnesota; ✱Bradley P. Carlin, The University of Minnesota, Division of Biostatistics MMC 303, 420 Delaware St. S.E., Minneapolis, MN 55455, *brad@biostat.umn.edu*

Key Words: areal data, continuous time data, CAR model, temporal gradient

The exact mechanisms relating exposure to ultraviolet (UV) radiation and elevated risk of skin cancer remain the subject of debate. For example, there is disagreement on whether the main risk factor is duration of the exposure, its intensity, or some combination of both. There is also uncertainty regarding the form of the dose-response curve, with many authors believing only exposures exceeding a given (but unknown) threshold are important. In this paper we estimate such thresholds using hierarchical spatial logistic models based on a cohort of x-ray technologists for whom we have self-reports of time spent in the sun and numbers of blistering sunburns in childhood. A secondary goal is to explore the temporal pattern of UV exposure and its gradient. Changes here would imply that identical exposure self-reports from different calendar years may correspond to differing cancer risks.

The Association Between Short-Term Exposure to Ozone and Risk of Mortality in the United States

✱Michelle L. Bell, Yale University, 205 Prospect Ave., New Haven, CT 06511, *michelle.bell@yale.edu*; Roger D. Peng, Johns Hopkins Bloomberg School of Public Health; Aidan McDermott, Johns Hopkins Bloomberg School of Public Health; Scott L. Zeger, Johns Hopkins Bloomberg School of Public Health; Jonathan M. Samet, Johns Hopkins Bloomberg School of Public Health; Francesca Dominici, Johns Hopkins University

Key Words: air pollution, human health, Bayesian modeling, epidemiology, mortality, ozone

We explored effects of ozone on mortality risk for 95 U.S. urban communities for the period 1987–2000. We applied Bayesian hierarchical models that first estimated associations within each community with single and distributed lag Poisson models, adjusted for weather, day of the week, and seasonality and long-term trend. In a second stage we combined community-specific estimates to generate national effects, accounting for the within-county and between-county statistical variances. Threshold effects were evaluated by estimating the nonlinear exposure-response in multiple model forms. A 10 ppb increase in the previous week’s ozone was associated with a 0.52% (95% posterior interval 0.27-0.77%) in mortality risk. Community-specific and national estimates were robust to adjustment by PM10 and PM2.5. Results from threshold models consistently show that effects persist at low ozone levels.

Health Impacts Under Heat Waves

✱Bo Li, National Center for Atmospheric Research, 1300 30th St. Apt. B3-24, Boulder, CO 80303, *boli@ucar.edu*; Doug Nychka, National Center of Atmospheric Research

It is well known that heat waves can cause a number of deaths and that it has been predicted future heat waves in Europe and North America will become more intense, more frequent, and longer-lasting in the second half of the 21st century. The deaths from past heat waves and the predicted growth of future heat waves raise an urgent question: What is the association between human health and heat waves? We tackle this question by studying the morbidity and weather data in five major Wisconsin cities. A generalized additive model (GAM) with space-time varying parameters is employed to analyze the relationship between the morbidity (hospital and emergency room admissions) and weather parameters in Wisconsin. This analysis is coupled with additional information from air pollution monitoring stations and adjusted by various confounding variables.

Health Effects of Chemical Constituents and Sources of Fine Particulate Matter

✱Francesca Dominici, Johns Hopkins University, Department of Biostatistics, 615 N Wolfe Street, 21230, Baltimore, MD 21205, *fdominic@jhsph.edu*

Key Words: particulate matter, chemical mixture, Bayesian hierarchical models

Optimal regulatory control of particulate matter (PM) is hindered by an insufficient understanding of the chemical characteristics of the PM mixture that determine its toxicity. Approximately 50 correlated chemical components of PM2.5 are monitored almost daily for several hundred monitoring stations in the US. A critical question is whether a small subset of these chemical constituents, alone or through their interactions, are responsible for the estimated harmful effects of PM. To estimate the toxicity of the PM mixture we have assembled a national data set on health outcomes, PM chemical constituents, and all the important confounders. In this talk, we will present Bayesian hierarchical regression models for analyses of these large and complex national data sets. Our modeling approaches will have the ultimate goal of estimating the health effects of PM size and chemical composition.

117 Variable Selection with High-Dimensional Data

WNAR, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, Section on Statistics in Epidemiology, IMS, Biometrics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

Challenge of Dimensionality in Model Selection and Classification

✱Jianqing Fan, Princeton University, Department of Operations Res and Fin Eng., Princeton, NJ 08544, *jqfan@princeton.edu*; Yingying Fan, Harvard University

Key Words: Classification, High-dimensionality, Variable Selection, FAIR

Model selection and classification using high-dimensional features arise in many statistical studies. The impact of dimensionality on classifications is poorly understood. We show that high-dimensional classification using all features can be as bad as the random guess and that almost all linear discriminants perform as bad as the random guess. Thus, it is important to select a subset of important features for high-dimensional classification, resulting in Features Annealed Independence Rules (FAIR). The connections with the sure independent screening (SIS) and iterative SIS of Fan and Lv (2007) in model selection will be elucidated and extended. The choice of the optimal number of features is proposed based on an upper bound of the classification error. Numerical studies support our theoretical results and demonstrate convincingly the advantage of our new classification procedure.

Variable Inclusion and Shrinkage Algorithms

✱Peter Radchenko, University of Southern California, Los Angeles, CA 90089, *peter.radchenko@marshall.usc.edu*; Gareth M. James, University of Southern California

Key Words: Variable Selection, Shrinkage, Lasso, Dantzig Selector

The Lasso is a popular and computationally efficient procedure for automatically performing both variable selection and coefficient shrinkage on linear regression models. One limitation of the Lasso is that the same tuning parameter is used for both variable selection and shrinkage. As a result, it typically ends up selecting a model with too many variables to prevent over shrinkage of the regression coefficients. We suggest a new class of methods called “Variable Inclusion and Shrinkage Algorithms” (VISA). Our approach is capable of selecting sparse models while avoiding over shrinkage problems and uses a path algorithm so is also computationally efficient. We show through extensive simulations that VISA significantly outperforms the Lasso and also provides improvements over more recent procedures, such as the Dantzig selector, Relaxed Lasso, and Adaptive Lasso.

Fast Regularization Paths

✱Trevor Hastie, Stanford University, Statistics Department, Sequoia Hall, Stanford, CA 94305, *hastie@stanford.edu*

Key Words: Lasso, Regularization Path, Coordinate Descent

The Lasso, Elastic Net and related techniques perform regularization and variable selection simultaneously. They can be applied in a variety of settings, such as the family of generalized linear models, multinomial regression and the Cox proportional hazards model. Typically we do not fit just one model, but a series of models as we vary the tuning parameter, creating a so-called regularization path. In this talk we discuss fast algorithms based on coordinate descent for fitting such paths, and modifications that can deal with very large sparse data sets.

118 Bayesian Methods for Diagnostic Test Meta-Analysis ●

Section on Health Policy Statistics, Section on Bayesian Statistical Science, Section on Statistics in Epidemiology

Monday, August 4, 10:30 a.m.–12:20 p.m.

Bayesian Meta-Analysis of Diagnostic Test Accuracy Studies: Recent Developments

Constantine Gatsonis, Brown University; ✱Carolyn Rutter, Group Health Cooperative, Center for Health Studies, 1730 Minor Ave., Ste 1600, Seattle, WA 98101, *rutter.c@ghc.org*

Key Words: Hierarchical Summary ROC, diagnostic accuracy, meta-analysis, hierarchical models, diagnostic tests

Interest in evidence-based diagnosis has grown rapidly in recent years and has highlighted the need for systematic reviews in this area. We will discuss statistical methods for diagnostic accuracy studies, with a focus on studies reporting estimates of sensitivity and specificity. The need to account for between-study differences in the threshold for test positivity is a fundamental aspect of systematic reviews of test accuracy has led to Summary Receiver Operating Characteristic curve analysis. The reviews also need to account for other sources of within- and between-study heterogeneity and to address issues such as errors in the reference standard, the use of multiple cutpoints, and verification bias. We will review and compare hierarchical and mixed model methods and will conclude with a discussion of continuing challenges in both statistical technique and reporting of results.

Bayesian Methods for Diagnostics: A Case Study of a Registration Trial

✱Scott M. Berry, Berry Consultants, 3145 Chaco Canyon Drive, College Station, TX 77845, *scott@berryconsultants.com*; Donald A. Berry, M.D. Anderson Cancer Center

A case study of a registration study for a diagnostic is presented. An adaptive sample size approach is used to monitor the sensitivity and specificity of the diagnostic. Bayesian analyses of the “true” sensitivity and specificity are discussed when the gold standard is measured with error.

Empirical Study of Bayesian Methods for Meta-Analyses of Diagnostic Test Data Using a Large Database of Studies in the Medical Literature

✱Christopher H. Schmid, Tufts Medical Center, 800 Washington St., Boston, MA 02111, *cschmid@tuftsmedicalcenter.org*

Key Words: Meta-analysis, Diagnostic tests, Bayesian, Meta-regression, ROC, Hierarchical model

Researchers apply a variety of statistical methods for meta-analysis of data from diagnostic test studies. Typically, each study provides a 2x2 table of true and false positives and negatives from which sensitivity and specificity are derived. Because of the bivariate structure of the data and the typically substantial between-study heterogeneity, Bayesian bivariate models are attractive. It has recently been shown that the bivariate normal model is equivalent to a special case of the hierarchical summary ROC (HSROC) model. We examine the performance of Bayesian forms of these models, contrasting them with simpler models of diagnostic odds ratios and univariate sensitivity and specificity in 265 meta-analyses of diagnostic tests from the medical literature. Special attention focuses on the effect of covariates in meta-regression models in attempting to understand the heterogeneity.

119 Estimating the Exposure to Risk ●▲

Section on Risk Analysis, Section on Statisticians in Defense and National Security, Section on Health Policy Statistics, Section on Statistics in Epidemiology

Monday, August 4, 10:30 a.m.–12:20 p.m.

A Perspective on Adversarial Risk Analysis

✱Nozer Singpurwalla, The George Washington University, 2140 Penn. Avenue, Washington, DC 20052, *nozer@gwu.edu*

Key Words: decision theory, game theory, decision trees

In classical risk analysis, one chooses actions that maximize ones expected utility against the actions of a benevolent adversary, namely, nature. In adversarial risk analysis, the adversary is an active opponent whose aim is to deny the decision maker from taking actions that maximize the decision maker's expected utility. Game theory has been advocated as the appropriate methodology for the treatment of adversarial risk. This point of view has been challenged by some decision theorists who claim that the methods of decision theory are adequate. In this expository talk, I highlight the issues raised in the context of adversarial risk and point out the circumstances under which decision theory and game theory are suitable methods.

Sampling Based on Two Types of Prior

✱Michael E. Tarter, University of California, Berkeley, School of Public Health, University Hall, University of California, Berkeley, CA 94720, *Tarter@berkeley.edu*; Ai-Chu Wu, Consultant

Key Words: Efficiency, Inverse Cumulative Distribution Function, Nested design, Numerical Integration, Orthogonal Polynomials, Robustness

According to Doull et al., a risk assessor often must know .".. (a) the toxicity profile of the agent of interest; (b) its interactions with living systems; and (c) the known or projected exposure scenarios: to whom, how much, by which route(s), and how often. In practice, however, complete information is seldom available. Nonetheless, decisions still must be made." Since information sources a, b, and c must often be combined, we studied sampling procedures designed to optimize assessment accuracy by using multiple sources of prior agricultural and environmental risk information. We combined systematic sampling based on Gaussian-Jacobi quadrature, on the one hand, and random sampling approaches that are often dealt with today by Neyman allocation strategies, on the other. In this way we utilized prior information about some strata's standard deviations, and some stratas' ranked means.

Evaluating the Risk and Exposure of Adverse Outcomes of Drug Products: Recent Experiences and Lessons

✱Mark Levenson, U.S. Food and Drug Administration, Bld. 22, Rm. 5245, 10903 New Hampshire Ave., Silver Spring, MD 20993, *mark.levenson@fda.hhs.gov*

Key Words: pharmacoepidemiology, risk, meta-analysis, drug safety

The "gold standard" to evaluate the efficacy of a drug is the randomized controlled trial. Trials provide valuable information on the safety as well. However, if an adverse event is rare or has a high background rate in the population, then a trial may not be adequate to evaluate the event risk. Additionally, safety issues may emerge after efficacy trials, and information may not be collected to measure the adverse event accurately. To overcome these limitations and to supplement information from trials, other data sources and approaches may be used. These include meta-analysis, observational studies, and adverse-event spontaneous reporting systems. This talk will review recent experiences with these approaches to evaluate drug risk, including the association of suicidality and antidepressants and the adverse outcomes associated with a cardiac-surgery anti-bleeding drug.

Exposure to Risk from Motor Vehicles and Other Consumer Products

✱Duane Steffey, Exponent, Inc., 149 Commonwealth Drive, Menlo Park, CA 94025, *dsteffey@exponent.com*

Key Words: design safety, injury data, field performance, recall

Estimating risk of injury from motor vehicles and other consumer products requires not only data on the occurrence of injuries but also some measure of the opportunity for injury (i.e., the exposure). Comparative assessments of motor vehicle risk often depend greatly on the measure of exposure used. Crash-based metrics more easily permit adjustment for confounding factors. Exposure-based metrics reflect better the influence of driver and environmental factors in accident scenarios. Data on exposure to general classes of consumer products may be found in customer surveys sponsored by industry groups or government agencies. For risk analyses of specific products, exposure data (i.e., the number of units at risk of failure) are typically obtained from a company's production, shipping, or sales records.

120 Planets Around Other Suns: Inference and Experimental Design for Exoplanet Studies ●▲

IMS, Section on Physical and Engineering Sciences, Section on Bayesian Statistical Science

Monday, August 4, 10:30 a.m.–12:20 p.m.

Detecting New Extrasolar Planets with a Bayesian MCMC Kepler Periodogram

✱Phil Gregory, University of British Columbia, Dept. of Physica & Astronomy, 6224 Agricultural Road, Vancouver, BC V6T 1Z1 Canada, *gregory@physics.ubc.ca*

Key Words: Extrasolar planets, Bayesian statistics, Markov chain Monte Carlo, Periodogram, time series

Are we on the threshold of discovering life elsewhere in the universe? With the mounting discoveries of extrasolar planets there is a sense that we are well underway. Astronomers searching for the small signals induced by planets inevitably face significant statistical challenges. This talk illustrates how a Bayesian reanalysis of published radial velocity data sets is providing strong evidence for additional planetary candidates. The Bayesian inference is accomplished with a parallel tempering Markov chain Monte Carlo (MCMC) algorithm which acts as a powerful multi-planet Kepler periodogram for both parameter estimation and model selection. The MCMC algorithm is embedded in a unique two-stage adaptive control system that automates the tuning of the proposal distribution parameters through an annealing operation.

Posterior-Guided Importance Sampling for Calculating Marginal Likelihoods, with Application to Bayesian Exoplanet Searches

✱Jim Crooks, Duke University, Dept. of Statistical Science, Box 90251, Durham, NC 27708-0251, *bigjim@samsi.info*

Key Words: Astronomy, extrasolar planets, Bayesian statistics, Importance sampling, Markov chain Monte Carlo, Experimental design

This talk presents a novel method of high-dimensional stochastic integration for calculating marginal likelihoods when a posterior MCMC sample is available. The accuracy and efficiency of importance sampling depends strongly on the importance function having approximately the same shape as the target function. In high dimensions, such importance functions can be extremely hard to build. However, in the Bayesian context, the target is proportional to the posterior, so a posterior MCMC sample can be used to guide the choice of importance function. We do this by centering multivariate student distributions around some of the MCMC points and then adapting their parameters to minimize the variance in the importance weights. Our work is done in the context of exoplanet detection, first in determining how many planets are present, and second in scheduling telescope time.

121 Adaptive Designs: Perspectives from Academia, Industry, and Regulatory ●▲

Biopharmaceutical Section, Biometrics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

Utility of Bayesian Methods in Early- and Late-Phase Adaptive Design Clinical Trials

✱Sue-Jane Wang, U.S. Food and Drug Administration, 10903 New Hampshire Ave, MailStop WO22, Room 6105, CDER, Silver Spring, MD 20993, *suejane.wang@fda.hhs.gov*

Key Words: A&WC, clinical trial, early adaptive trial design, type II error

The recent advances in adaptive design methodology for evaluation of an experimental treatment extend from sample size re-estimation based on accumulating interim data to mid-stream adaptation of primary efficacy endpoint, treatment dose/schedule/duration, or responsive patient population in either early and/or late phases controlled trials. There have been various enthusiastic attempts to bring in Bayesian tools in the design stage, the analysis stage, and the inference stage. In this paper presentation, the role and utility of Bayesian methods for early phase drug development versus late phase drug development will be elucidated using typical and novel scenarios. The contrast between decision issues versus inference issues will be highlighted. Implementation issues related to practical usefulness and interpretability will be discussed.

Population Enrichment Within a Group Sequential Design

✱Cyrus R. Mehta, Cytel Inc., 675 Massachusetts Avenue, Cambridge, MA 02139, *mehta@cytel.com*; Ping Gao, The Medicines Company; James Ware, Harvard University

Key Words: Enrichment Design, Subgroup Selection, Adaptive Design, Cardiology Trials, Group Sequential, Sample Size Increase

We present a method for combining group sequential stopping rules with population enrichment within the framework of a confirmatory clinical trial. The approach is motivated by the need for greater efficiency in large cardiology trials where event rates are low and efficacy gains relative to the current standard of care are likely to be small. Since such trials typically involve sample sizes in the thousands, group sequential designs with early stopping for benefit or futility are commonly adopted. These designs can be improved if patient eligibility is initially open to a broad population with the option, at an interim analysis, to restrict future enrollment only to prespecified subgroups that appear to be benefiting from the new treatment, possibly accompanied by a data dependent sample size increase.

Adaptive Designs: Why, How, and When?

✱Christopher Jennison, University of Bath, Department of Mathematical Sciences, Bath, International BA15 1EP United Kingdom, *C.Jennison@bath.ac.uk*

Key Words: Clinical trials, Adaptive designs, Group sequential tests, Combination tests, Seamless transition

The applications of adaptive design methodology are diverse and one should not expect a simple answer to when it may be beneficial to use such a design. We shall consider a range of application areas and offer conclusions for each. Conventional group sequential tests are relevant to problems of sample size reassessment. However, adaptive designs offer a flexible response to unexpected changes of intent or to rescue an inadequately planned trial. Combination tests at the heart of many adaptive designs play a key role in two-stage studies set up to include treatment selection, change of endpoint, or change of focus to a subpopulation. Combining two of the traditional phases of clinical testing with a “seamless transition” may have benefits of continuity and, ultimately, more rapid recognition of an effective treatment—as long as logistical issues can be addressed.

122 The Role of Priors in Bayesian Analysis of Complex, Real-World Problems ●▲

International Society of Bayesian Analysis, Section on Bayesian Statistical Science, Section on Health Policy Statistics

Monday, August 4, 10:30 a.m.–12:20 p.m.

Combining Experts in Prior Elicitation

✱Judith Rousseau, University Paris Dauphine, Place du marechal deLattre de Tassigny, Paris, International 75 016 France, *rousseau@ceremade.dauphine.fr*; Guihenneuc Chantal, University Paris 5; Albert Isabelle, INRA; Donnet Sophie, University Paris Dauphine; Mengersen Kerrie, Queensland University of Technology; Lowchoy Samantha, Queensland University of Technology

Key Words: Bayesian statistics, prior elicitation, hierarchical modelling

In many applications there is a need to use expert knowledge to construct prior distributions in a Bayesian approach. There is a vast literature on the elicitation of prior distributions using expert knowledge. However there remains some open questions. In particular when more than one experts are interviewed and we want to combine mathematically their opinions the approaches the are proposed in the literature are not entirely satisfactory since they do not really take into account the possible dependence between the experts. In this paper we propose to use a hierarchical modeling to combine the experts’ opinions into a single prior. We apply this method in two different setups : a dose-response model as is typically used in food risk analysis and a sociological model. A comparison is made with other methods.

Bayesian Modeling of Neuron Death

Anthony Pettitt, Queensland Univeristy of Technology; ✱Gareth Ridall, Lancaster University, Maths and Stats, Lancaster, LA1 4YW United Kingdom, *g.ridall@lancs.ac.uk*

Key Words: neuron death, model choice, informative prior, Bayes factor, motor neuron disease

Amyotrophic lateral sclerosis (ALS), a type of motor neuron disease, is a degenerative neurological disease. Mathematical models have been put forward to explain the death process for the numbers of surviving nerve cells. For ALS Ridall et al (Appl Statistics, 2007) developed a Bayesian model to allow clinicians to assess disease progress by estimating a patient's remaining number of motor units by analyzing data measured from a given muscle in response to electrical stimulation of a nerve. In that approach many informative priors were used but a flat prior was used for the number of units. Data from previous studies carried out with the patient were considered separately. We investigate the use of a two stage process to incorporate previous patient data. We use the Bayes factor to distinguish between competing models of neuron death which have different biological interpretations.

Hierarchical Models for Combining Data from Multiple Sources for Risk Assessment

Louise Ryan, Harvard School of Public Health; ✱Tianxi Cai, Harvard University, Department of Biostatistics, 655 Huntington Ave., Boston, MA 02115, *tcai@hsph.harvard.edu*

Assessing the risk of rare adverse events associated with environmental or pharmaceutical exposures generally encounters the challenge of dealing with low-powered studies. Power to quantify such risks can be boosted by synthesizing data from multiple studies. While Bayesian hierarchical models provide a natural tool in such settings, it can be a challenge to construct appropriate prior distributions, and the results are often sensitive to the assumptions being made. We present several case studies that illustrate these challenges, along with effective solutions.

123 Helping Data Users Better Understand the American Community Survey ●▲

Section on Government Statistics, Social Statistics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

Helping Data Users Better Understand the American Community Survey

✱Linda A. Jacobsen, Population Reference Bureau, 1875 Connecticut Ave., NW, Suite 520, Washington, DC 20009, *ljacobsen@prb.org*; ✱Peter Lobo, NYC Department of City Planning, 22 Reade Street - 4W, New York, NY 10007-1216, *plobo@planning.nyc.gov*; ✱Ken Hodges, Claritas Inc., 53 Brown Road, Ithaca, NY 14850, *khodges@claritas.com*; ✱John Thompson, National Opinion Research Center, *Thompson-John@norc.org*; ✱Joseph Salvo, NYC Department of Planning; ✱Kennon Copeland, National Opinion Research Center

As a part of the 2010 Census re-engineering, the ACS replaced the traditional decennial census long form, allowing for up-to-date information every year, rather than just once a decade. Beginning in 2006, the Census Bureau released single-year estimates for the nation and for states, counties, and other geographic areas with populations of 65,000 and greater. In 2008, the first set of multiyear period estimates will be released for data collected between 2005 and 2007. These 3-year period estimates will include geographic areas with populations of 20,000 and greater. To prepare for this release, the Census Bureau is working with a set of experts to develop training and informational materials geared toward specific audiences. This session will present the results of this effort and focus on the general public, state and local government, the business community, and the advanced ACS user.

124 Directing a Biostatistics Core in Biomedical Research ●▲

Council of Chapters, Biometrics Section, Biopharmaceutical Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

Directing a Biostatistics Core in Biomedical

✱Michael Kutner, Emory University, Department of Biostatistics, Atlanta, GA 30322, *mkutner@sph.emory.edu*; ✱Karen Bandeen-Roche, Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, 615 North Wolfe Street, Baltimore, MD 21205-2179, *kbandeen@jhsph.edu*; ✱Joseph Hogan, Brown University, Center for Statistical Sciences, 121 South Main Street, Box G-S121-7, Providence, RI 02912, *jwh@brown.edu*; ✱KyungMann Kim, University of Wisconsin-Madison, Department of Biostatistics & Medical Informatics, Madison, WI 53792, *kmkim@biostat.wisc.edu*; ✱Susan Ellenberg, University of Pennsylvania, Dept of Biostatistics and Epidemiology, Philadelphia, PA , *sellenbe@mail.med.upenn.edu*

Key Words: Biostatistics, Consulting

A common challenge in the practice of biostatistics entails dealing with broader scientific and operational issues. As such, the goal for this invited panel session is to provide a venue for the exchange of ideas concerning the challenges and solutions that arise in serving as director of either a biostatistic core or biostatistics shared resource within a university-based research center. The panel is composed of biostatisticians who have extensive experience serving in such a capacity. Both statistical and nonstatistical aspects of core direction will be discussed.

125 Communicating Statistics Through Collaborative Problemsolving ●▲

Section on Statistical Consulting, Committee on Career Development, Biopharmaceutical Section, Section on Quality and Productivity, Section on Teaching Statistics in the Health Sciences

Monday, August 4, 10:30 a.m.–12:20 p.m.

Our Collaborative and Value-Creating Problemsolving Process

✱Arnold Goodman, Collaborative Data Solutions, 92861, *thegoodman@dslextreme.com*

Key Words: problem solving, value creation, stakeholder orientation, collaboration

Since early 1900, there has been a continuing evolution in problems, data sets, technology, and methodology, plus the number and perspective of stakeholders. Significant problems today are highly complex in structure and multidisciplinary in nature, with massive data sets. In response, we developed a value-creating process to solve stakeholder problems, view of collaboration's success factors, and fabric of collaboration success factors woven into the process. It expands Fisher's 3 problems in reducing data and Mallow's 0th problem. Five factors critical to collaboration are community of collaborator relationships, commitment to community goals, communication in community language, change produced by contributions to community goals, and control from stakeholder evaluation. We know no other process as problem and data driven, stakeholder focused, value-oriented, or collaborative as ours.

Collaborative Problemsolving Case Studies

✱Janet Myhre, Mathematical Analysis Research Corp., 4239 Via Padova, Claremont, CA 91711, *MARCmath@aol.com*

Key Words: problem solving, value creation, stakeholder orientation, collaboration, illustrative case studies

This presentation complements "Our Collaborative and Value-Creating Problemsolving Process with Case Studies," which illustrate its aspects. Case studies exemplify the importance of collaboratively: (1) defining the stakeholders' actual problem needs, (2) specifying the data and theory resources required to satisfy them, (3) designing the analysis and software tools to produce a problem solution, (4) building the problem solution, and finally (5) creating value for the stakeholders from this solution. In addition, they will portray what collaboration is and how to do it successfully. Case studies include 'Combination of Data Sources for Estimation of Success or Failure,' 'Pharmaceutical Product Development Projections,' 'Development of Event Tree Software for Safety,' and 'Forecast Models for the Salinity of the Salton Sea.'

Teaching Consulting and Collaboration Skills in a Graduate Statistics Program

✱Daniel R. Jeske, University of California, Riverside, Department of Statistics, Room 2605 STAT-COMP Building, Riverside, CA 92521, *daniel.jeske@ucr.edu*

Key Words: Cat Food Recipes, Bradley-Terry Models, Statistical Consulting, Graduate Student Education

It is shown how student participation in a real consulting project can be leveraged to achieve the dual goals of (i) developing statistical consulting skills in graduate students, and (ii) enhancing the instructional effectiveness of statistical methodology. Achieving these goals is the primary mission of the Statistical Consulting Collaboratory at the University of California, Riverside. The paper gives a detailed illustration of the how the goals were achieved by reporting on an interesting case study, with special emphasis given to describing the involvement of students and the alternative ways in which the project found its way into classrooms. (This is joint work with Scott M. Lesch and Hongjie Deng.)

Human Challenges to the Collaborative and Value-Creating Problemsolving Process

✱Doug Zahn, Zahn and Associates, 249 Timberlane Road, Tallahassee, FL 32312, *zahn@stat.fsu.edu*

Key Words: relationship, wanted conversation, breakdown, meeting, attitude, feelings

Many challenges to the collaborative and value-creating problem solving process come from the fact that one is doing this activity with other humans: creating synergistic relationships; sorting out what is to be done–rather than succumbing to the cry, "We have talked long enough! Let's get to work!"; recovering from the breakdowns that inevitably occur in the process; and having effective meetings. I will explore strategies for sorting out these challenges and dealing with them effectively.

126 Integrating Statistics and Bioinformatics Curricula ●

Section on Teaching Statistics in the Health Sciences, Biopharmaceutical Section, Section on Statistical Education

Monday, August 4, 10:30 a.m.–12:20 p.m.

Applying the GAISE Framework to a Bioinformatics Course

✱Monnie McGee, Southern Methodist University, SMU P.O. Box 750332, 3225 Daniel Ave Room 144 Heroy, Dallas, TX 75275-0332, *mmcgee@smu.edu*

Key Words: GAISE Guidelines, Education, Bioinformatics, Classroom Technology, Software

The Guidelines for Assessment and Instruction in Statistics Education (GAISE) Project was completed in 2005. The guidelines were meant specifically for introductory statistics courses, but they can apply to other courses, also. Recommendations from this project include emphasize statistical literacy and develop statistical thinking, use real data; stress conceptual understanding rather than mere knowledge of procedures, foster active learning in the classroom, use technology for developing conceptual understanding and analyzing data, and use assessments to improve and evaluate student learning. In my talk, I will discuss how each of these guidelines can be directly applied or modified to form the framework for a course in bioinformatics. I will also discuss other guidelines that are applicable to bioinformatics courses.

Introducing Bioinformatics as Part of Statistics Curriculum at Purdue

✱Olga Vitek, Purdue University, IN 47906, *ovitek@stat.purdue.edu*

Key Words: bioinformatics, education, microarrays

Problems in molecular biology offer exciting opportunities in statistical research, but require specialized interdisciplinary knowledge and skills. Statistics students with little biology background are sometimes reluctant to enter the field, but can be encouraged if they are exposed to it early in the program. Our department offered an introductory course that can be taken during the first two years in graduate school. The course focuses on the analysis of gene expression microarrays, and builds upon a series of successful workshops offered by the bioconductor community. The course emphasizes statistical methods in the "large p small n" paradigm, inference in the context of biological knowledge and annotations, and computing with R and bioconductor. We found that the course helped stimulate the interest in this area, and motivated a number of students to pursue research in this field.

Strategies for Integrating Biological, Computer, and Statistical Knowledge in Teaching Bioinformatics

✱Paul Schliekelman, The University of Georgia, Department of Statistics, 204 Statistics Building, Athens, GA 30602-1952, *pdschlie@stat.uga.edu*

Key Words: bioinformatics, teaching, curriculum

One of the major challenges in teaching bioinformatics is the broad skill set required of students. A successful bioinformaticist requires genuine biological knowledge, good programming skills, and strong math/statistical skills. It is difficult to teach even introductory courses without some level of knowledge in all three of these areas and very few students come adequately prepared. I will discuss the evolution of my strategies in five years of teaching a two semester statistical bioinformatics course. I will also discuss the approach to this issue in formulating a curriculum for a new PhD program in bioinformatics at the University of Georgia and the results after one year with our first batch of graduate students.

The Flip Side: Teaching Statistics to Bioinformatics Students

✱Mark Segal, University of California, San Francisco, 185 Berry Street, Lobby 4, Suite 5700, San Francisco, CA 94107, *mark@biostat.ucsf.edu*

Key Words: Bioinformatics, Computational Biology, Case Studies, Doctoral Program

At UCSF there is a doctoral program in†Bioinformatics but none in (bio) statistics. Accordingly, faculty from the Division of Biostatistics†are involved in teaching and mentoring†Bioinformatics students in various aspects of data analysis.†One standing course offering consists of a “case studies”† lecture series, wherein illustrative examples from the computational biology literature that showcase†applications (both good and bad)†of particular†statistical methodologies are explained and dissected. This format also allows faculty to introduce the own research, with the goal of fostering dissertation and/or project collaborations. Strengths and weaknesses observed to date are discussed.

127 Census Coverage Measurement ▲

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

Relaxing the Autonomous Independence Assumption for Census Coverage Measurement Dual System Estimates

✱Richard A. Griffin, U.S. Census Bureau, 104 Arrowhead Circle, Lansdale, PA 19446, *richard.a.griffin@census.gov*

Key Words: Dual System Estimation, Taylor Linearization, Sensitivity Analysis

This paper develops expressions for the expected value and variance of the dual system estimate relaxing the autonomous independence assumption. The development assumes all members of the population are exposed to possible inclusion in both the coverage measurement survey and the census. Taylor Linearization is used as in Wolter’s 1986 *JASA* paper; however, the captures of persons in the same housing unit are no longer assumed to be independent. A sensitivity analysis is done to analyze the effect on the bias and multinomial model variance of the dual system estimate for varying amounts of capture dependence. Data from the Census 2000 Accuracy and Coverage Evaluation Survey is used to estimate potential levels of capture dependence for persons in the same housing unit.

Using Continuous Variables as Modeling Covariates for Net Coverage Estimation

✱Vincent T. Mule, U.S. Census Bureau, 15007 Nebraska La, Bowie, MD 20716, *vincent.t.mule.jr@census.gov*; Don Malec, U.S. Census Bureau; Jerry Maples, U.S. Census Bureau; Teresa Schellhamer, U.S. Census Bureau

Key Words: Dual System Estimation, Logistic Regression

The Census Bureau uses dual system estimation as one method to evaluate the coverage of the decennial census. This estimation for previous censuses has used post-stratification to minimize the impact of correlation bias on the population estimates. For 2010, we are planning on using logistic regression modeling instead of post-stratification cells. Logistic regression gives us the option of using variables in the model as continuous variables instead of having to form groupings. This research present our initial results of the impact of using certain continuous variables in the model development and the resulting population estimate.

Direct Estimates as a Diagnostic for Dual System Estimators Based on Logistic Regression

✱Mary H. Mulry, U.S. Census Bureau, 2884 Manorwood Trail, Fort Worth, TX 76109, *mary.h.mulry@census.gov*; Bruce D. Spencer, Northwestern University; Tom Mule, U.S. Census Bureau; Nganha Nguyen, U.S. Census Bureau; Eric Schindler, U.S. Census Bureau

Key Words: census, Accuracy and Coverage Evaluation Survey, 2010 Census Coverage Measurement Program

The 2010 Census Coverage Measurement Program (CCM) is preparing to use logistic regression modeling in the estimation of net census coverage error rather than post-stratification, the approach used for previous censuses. The most important objective for the CCM is to obtain separate estimates of erroneous census inclusions and census omissions. The plan for estimating census omissions is to sum estimates of net coverage error and erroneous enumerations. The net error estimates will be based on dual system estimation formed with separate logistic regression models for the correct enumeration rate and the match rate. Direct estimates at the block cluster level aid in variable selection by comparing the accuracy of estimates based on logistic regression models (or post-stratification designs) with and without a variable for groups of the clusters with different characteristics.

Assessing Synthetic Error via Markov Chain Monte Carlo Techniques

✱Andrew Keller, U.S. Census Bureau, 4600 Silver Hill Rd., Room 4H260D, Washington, DC 20233, *andrew.d.keller@census.gov*

Many small domain estimates require a precise, direct estimate of the within domain variability as one component. However, due to small sample size, the precision of within small domain direct variance estimates of census coverage is questionable. In this paper, Markov chain Monte Carlo (MCMC) techniques are applied to develop a model-based estimate of the within domain variability as part of the estimation process. For this particular application, variability within state is modeled via a random effects model where the census block is the replicate. Ultimately, this block-level model is applied to evaluate synthetic error of the small domain.

Demographic Analysis in 2008 and Beyond

✱Dean Judson, U.S. Census Bureau, 15000 Fort trail, Accokeek, MD 20607, *dean.h.judson@census.gov*; J. Gregory Robinson, U.S. Census Bureau; Kirsten K. West, U.S. Census Bureau; Antonio Bruce, U.S. Census Bureau; Bashiruddin Ahmed, U.S. Census Bureau; Frederick W. Hollman, U.S. Census Bureau

Key Words: coverage measurement, dual system, demographic analysis, demographic benchmarks, decennial census

The role of Demographic Analysis (DA) is to provide a coverage measurement estimate of decennial results independent from the (survey based) dual system estimator. In 2000, demographic analysis proved its worth. However, there are many improvements that should be considered. There are two major foci of interest: Constructing national DA measures and subnational demographic “benchmarks,” and evaluating components of DA measures. These research foci have proven important given the success of DA in 2000 and the need for continual improvement. This paper will discuss the potential of several new research topics in each focus area, and discuss their feasibility within the operational limits of the 2008 dress rehearsal and Census 2010.

128 Threshold Regression and Alternative Time Scales

Biometrics Section, Section on Nonparametric Statistics

Monday, August 4, 10:30 a.m.–12:20 p.m.

Analyzing Longitudinal Survival Data Using Threshold Regression

✱George A. Whitmore, McGill University, 184 Norwood Ave, Point Claire, QC H9R1R4 Canada, *george.whitmore@mcgill.ca*; Mei-Ling T. Lee, The Ohio State University

Key Words: non proportional hazards, proportional hazards, first hitting time, longitudinal data, passage time, Wiener process

Longitudinal survival data pose an interesting challenge for statistical modeling and inference, especially where the data also require a regression structure. The methodology of threshold regression is based on the concept that people's health status follows a latent stochastic process and event occurs when the health status process first reaches a threshold (a first hitting time). Encounters with this type of data structure abound in practical survival settings and there is a pressing need for simple regression methods to handle the longitudinal aspect and covariate data. This study looks at the formal conditions that must hold for this so-called uncoupling procedure to be valid. The conditions are examined in terms of both theory and practical application. The uncoupling procedure modifies the time scale for the analysis and can be used in conjunction with an operational time scale.

First-Hitting Time Models with Current Status Data

✱Xin He, The Ohio State University, B-116 Starling-Loving Hall, College of Public Health, 320 West 10th Avenue, Columbus, OH 43210, *xhe@cph.osu.edu*; Xingwei Tong, Beijing Normal University; Jianguo (Tony) Sun, University of Missouri

Key Words: Bivariate Wiener process, Current status data, First hitting time model, Joint analysis

We consider the analysis of failure time data where a failure is defined as the time at which a particular latent process crosses a threshold level. Several models, usually referred to as first-hitting time or threshold models, for this type of data are available. In particular, Whitmore et al. (1998) proposed such a model based on a bivariate Wiener process that models both the latent process and a marker process that is related to the latent process jointly. One advantage of the Wiener process model is that it can handle the situation where the effect of the marker process on the latent process is not proportional. We consider the analysis of this model when only current status data are available and the maximum likelihood estimation approach is developed. The approach is evaluated by simulation studies and applied to a motivating example about liver tumors in mice.

A Regression-Spline-Based Threshold Regression Model

✱Zhangsheng Yu, The Ohio State Univesity, College of Public Health, OH 43210, *zyu@cph.osu.edu*; Mei-Ling T. Lee, The Ohio State University

Key Words: First hitting time model, survival analysis, cross validation

For censored survival data, we investigate a threshold regression model with a semi-parametric covariate function. In estimating the nonparametric covariate function, the regression spline method are discussed. A cross validation method is proposed for the selection of tuning parameters. Simulation results show that the proposed estimates for both the nonparametric function and parametric coefficients are close to true values. The proposed variance estimates also perform very well. We apply regression spline based procedure to the western Kenya parasitemia study data. The proposed model provides a useful regression model for data sets where proportional hazards assumption is not true.

Bayesian Random Effects Threshold Regression with Application to Survival Data with Nonproportional Hazards

✱Michael Pennell, The Ohio State University, 320 West 10th Avenue, Columbus, OH 43210, *mpennell@cph.osu.edu*; George A. Whitmore, McGill University

Key Words: Bayesian methodology, Nonproportional hazards, Random effects, Survival analysis, Threshold regression, Wiener process

In biomedical studies, time to event data often violate the assumptions of Cox regression due to the presence of time-dependent covariate effects and unmeasured risk factors. An alternative approach is to model a subject's health status as a latent stochastic process that fails when it reaches a threshold value. Although these threshold regression models do not require proportional hazards, existing methods do not accommodate unmeasured covariates. To address this issue, we propose a Bayesian methodology which models the latent health status of an individual as a Wiener diffusion process with subject-specific initial state and drift. Posterior inference proceeds via an MCMC methodology with data augmentation steps to sample the final health status of censored observations. We demonstrate our method using mortality data from melanoma patients with nonproportional hazards.

129 Time Series II: Model Optimization and Evaluation

Business and Economics Statistics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

A Modification to Khandakar and Hyndman's ARIMA Model Selection Algorithm Using an Empirical Information Criterion

✱Brian C. Monsell, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *brian.c.monsell@census.gov*

Key Words: Automatic ARIMA model selection, out-of-sample forecast error, outlier regressors, calendar regressors, history diagnostics

Khandakar and Hyndman (2007) propose an automatic ARIMA model selection algorithm based on the Linear Empirical Information Criterion (LEIC), which penalizes the maximized log likelihood by a linear function of the number of parameters in the model. A penalty term is derived using an out-of-sample forecast performance criterion evaluated for a set of related time series. A modification to this algorithm will be examined that (a) includes selection of trading day, holiday and outlier regressors identified initially using a default model, and (b) incorporates 12-step ahead out-of-sample forecast performance. This modified algorithm will be applied to a set of Census Bureau series examined in McDonald-Johnson, Hood, Monsell and Li (2007). The model choices made by the algorithm will be compared to those made by Version 0.3 of X-12-ARIMA.

How Frequency of Measurement Affects Petroleum Product Price Pass-Through Models

✱Carol J. Blumberg, Energy Information Administration, 2035 Lyttonsville Road, Silver Spring, MD 20910-2264, *carol.blumberg@eia.doe.gov*; John Zyren, Energy Information Administration

Key Words: Petroleum, pass-through, asymmetry, time-series, temporal disaggregation, quadratic matching

The Energy Information Administration (EIA) has access to spot prices of conventional and reformulated gasoline (or RBOB), diesel fuel and other petroleum products on a daily basis. EIA surveys retail prices for these products on a weekly basis. Further, it collects data on volumes sold on a monthly basis. Since EIA reports prices by regions and formulation, when aggregated, the prices must be volume-weighted to get regional prices that are useful in modeling. Thus, when building models for predicting retail prices from spot prices, data frequency becomes an issue. This paper studies the impact of using different volume frequencies on the modeling of weekly price pass-through. The use of (i) an average volume (i.e., constant weight), (ii) the actual monthly reported volumes, and (iii) weekly estimates derived using quadratic-matching will be compared and contrasted.

Theoretical and Real Trading-Day Frequencies

✱Dominique Ladiray, INSEE, 18 boulevard Adolphe Pinard, Paris, 75014 France, *dladiray@yahoo.fr*

Key Words: Seasonal Adjustment, Tradin-Day Effect, X-12-ARIMA, Spectral Analysis

Most of economic indicators are linked to a daily activity which is usually summed up each month or each quarter. In this case, the number of working days, which varies from one month to another in a quasi-predetermined way, can explain some short-term movements in the time series. Due to the regularity of the calendar, this Trading-Day effect (TD) corresponds to specific frequencies and could be detected in the spectrum of the series. X-12-Arima implements such a "visual spectral test." In a first part, the "theoretical" TD frequencies are derived from the spectrum of the Gregorian calendar. But a real TD effect, a linear combination of these basic frequencies, can correspond to different frequencies. In a second part, the study of several thousands of real time series permits to exhibit the "real" TD frequencies, those one should use in a spectral test to detect a trading-day effect.

The Influences of Bridging Days, School Holidays, and Weather on German Time Series

✱Robert Kirchner, Deutsche Bundesbank, Wilhelm-Epstein-Str. 14, Frankfurt am Main, 60431 Germany, *robert.kirchner@bundesbank.de*

Key Words: seasonal adjustment, calendar adjustment, calendar variables

Traditionally, seasonal adjustment also involves the elimination of average calendar influences. By contrast, no separate adjustment usually takes place for bridging days, school holidays and weather-induced effects. These effects are investigated in the paper based on RegARIMA models. It will be shown that the three variables have a significant influence on the German production index. However, problems are apparent, these being that the adjusted results are not always plausible, the magnitude of the influences may depend on the business cycle and catch-up effects are probable but impossible to quantify. Hence, it is argued that the decision about which variables should be used for the calendar adjustment of officially published data should take into account more criteria than only statistical significance.

Further Results on Seasonal Time Series

✱David A. Dickey, North Carolina State University, Box 8203, NC State University, Raleigh, NC 27695-8203, *dickey@stat.ncsu.edu*

Key Words: nonstationarity, unit roots, asymptotics

Seasonal differencing is often applied when reporting, for example, monthly sales. new car sales are often reported as up or down from the same period last year. We modify the seasonal tests of Dickey, Hasza, and Fuller (1984) to investigate results for less typical, long period cases such as 1,440 minutes per day, 52 weeks per year, 365 days per year and so forth, getting some nice properties including a surprising effect of deterministic terms in the models. Seasonal cointegration has grown in popularity as well over the last several years and we compare several approaches to this kind of model for quarterly data. we focus on results for various specific alternatives suggested by the models in this class, comparing some standard models to one that is directed against the specific alternative under study.

130 Complex Hierarchical Bayesian Biostatistics ●▲

Section on Bayesian Statistical Science

Monday, August 4, 10:30 a.m.–12:20 p.m.

Analyzing Phylogenetic Recombination Data with Ancestral Recombination Graphs

✱Erik W. Bloomquist, University of California, Los Angeles, UCLA School of Public Health, Department of Biostatistics, Los Angeles, CA 90095, *bloomquist@gmail.com*; Marc A. Suchard, University of California, Los Angeles

Key Words: Phylogenetics, Recombination, Ancestral Recombination Graph, Coalescent, HIV

In the past ten years, advances in computing technology and Bayesian statistics have revolutionized the field of statistical phylogenetics. By allowing for more complex and realistic models, statistical phylogenetics can now more adequately model and analyze genetic datasets. In particular, Bayesian phylogenetic models have better handled the problem of putative recombination inference and detection. In this talk, I will review some of our group's research on Bayesian phylogenetic models for recombination. Afterward, I will discuss and present some of our current work regarding Ancestral Recombination Graphs.

Multivariate Bayesian Variable Selection and Voxel Selection in fMRI

✱Bradley McEvoy, University of California, Los Angeles, 1001 west stevens avenue #415, Santa Ana, CA 92707, *bmcevoy@ucla.edu*; Rajesh Nandy, University of California, Los Angeles

Key Words: fMRI, Bayesian, variable selection, multivariate

In fMRI, one would assume multivariate statistical techniques (e.g., multivariate regression) are more efficient than their univariate counterparts in dealing with the complex spatial correlation structure. The experimental paradigm's effect can be assessed over numerous voxels and does not require spatial smoothing of the data. However, multivariate methods suffer from bleeding artifacts induced by the p-value assignment schemes that inflate the size of the activated regions. In this talk, I adapt the multivariate Bayesian variable selection technique to resolve this issue. Further, this framework affords the following features; anatomical information, expert opinion of where activation is likely to occur and a thresholding mechanism can enter the statistical model.

Hierarchical Dynamic Time-to-Event Models for Post-Treatment Preventive Care Data on Breast Cancer Survivors

✱Freda Cooner, U.S. Food and Drug Administration, 1620 E JEFFERSON ST, APT 10, Rockville, MD 20852, *freda.cooner@fda.hhs.gov*; Sudipto Banerjee, The University of Minnesota

Key Words: Cure rate models, Dynamic survival models, Hierarchical models, Latent activation schemes, Markov chain Monte Carlo, Piecewise exponential distribution

This article considers modeling data in post-treatment preventive care settings, where cancer patients who have undergone the treatment discontinue seeking preventive care services. Clinicians and public health researchers are interested in explaining such behavioral patterns by modeling the time to receiving care while accounting for several patient and treatment attributes. Cure models are often preferred for data where a significant part of the population never experienced the endpoint. Building upon recent work on hierarchical cure model framework we propose modeling a sequence of latent events with a piecewise exponential distribution that remedies oversmoothing encountered in existing models with different latent distributions. We investigate simultaneous regression on the cure fraction and the latent event distribution and derive a flexible class of semiparametric cure models.

Bayesian Multivariate Longitudinal Models for Self-Reported Sex Behavior Data

✱Lily Altstein, University of California, Los Angeles, Los Angeles, CA , *laltstein@ucla.edu*; Robert E. Weiss, University of California, Los Angeles

Key Words: Bayesian Data Analysis, Random Effects Models, Multivariate Categorical Data, HIV Transmission-Prevention

HIV researchers regularly elicit detailed information on sexual behaviors in HIV+ and HIV-at-risk populations to assess counts of transmission behaviors. The CLEAR study randomized study participants either to an immediate else delayed counseling intervention to reduce their HIV transmission behaviors. We develop a joint multivariate hierarchical longitudinal model to study the interrelationships between the number of sexual partners and disclosure of participant's HIV-status to sex partners across several different partner types. The model analyzes the correlation between outcomes through both fixed and random effects and is a template for more complex analyses that would include number of sexual acts and condom use.

Model Selection for Clustered Outcome Common Predictor Effect Models

✱Juan Jia, University of California, Los Angeles, Department of Biostatistics, UCLA School of Public Health, Los Angeles, CA 90095, *jiajuan@ucla.edu*; Robert E. Weiss, University of California, Los Angeles

Key Words: Multivariate longitudinal data, Common effects, Leaps and bounds

Model selection is the problem of identifying the best models among a large number of competing models. Many model selection studies focus on variable selection. The Clustered Outcome common Predictor Effect (COPE) model clusters outcomes that respond in like fashion to all predictors. As the number of outcomes increases, the number of ways to cluster the outcomes increases rapidly. A leaps and bounds style algorithm is developed to efficiently seek a set of best COPE models. We apply the proposed method to multivariate longitudinal data from children of HIV+ parents. We seek the best ways to cluster nine psychometric sub-scales of the Basic Symptom Inventory and we use age, time in study, gender, season and parental drug use as predictors.

131 Theory and Applications of Curve Estimation ●▲

Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, IMS

Monday, August 4, 10:30 a.m.–12:20 p.m.

Nonparametric Methods in Gradual Change-Point Detection

✱Michael Baron, University of Texas at Dallas, Department of Mathematical Sciences, Richardson, TX 75083-0688, *mbaron@utdallas.edu*

Key Words: change-point, detection, delay, false alarm, gradual change, stopping time

Gradual change-point detection refers to the range of problems where the distribution of data slowly moves away from its stable, control state. Gradual disorders are difficult to detect immediately, however, it is often important to act quickly on them. Such situations occur in quality control, economics, epidemiology, public surveillance, and other types of process monitoring. In this talk, we focus on new nonparametric algorithms for detecting gradual changes. Unlike the traditional abrupt change-point problem, consistent estimators of the disorder time may be available. When little is known about the pre-change and especially, the post-change distribution of data, existing prior information on the potential disorder times cannot be neglected. Thus, special attention is paid to the development of Bayesian methods.

A New Semiparametric Procedure for Matched Case-Control Studies with Missing Covariates

✱Suojin Wang, Texas A&M University, Department of Statistics, Texas A&M University, College Station, TX 77843, *sjwang@stat.tamu.edu*; Samiran Sinha, Texas A&M University

Key Words: Conditional logistic regression, Estimating equation, Kernel regression, Missing at random

In this talk, we consider an easy-to-use semiparametric method for analyzing matched case-control data when one of the covariates of interest is partially missing. Missing covariate information in matched case-control study may create bias and reduce efficiency of the parameter estimates. In order to cope with this situation we propose a robust approach which is comprised of estimating some functionals of the distribution of the partially missing covariate using a kernel regression technique in a conditional likelihood framework. The large sample properties of the proposed estimator are investigated and the asymptotic normality is obtained. A simulation study is carried out to assess the performance of the proposed method in terms of robustness and efficiency. The proposed method is also applied to a real data set that motivates this work.

Confidence Bands Based on the Smoothing Spline

✱Vincent N. LaRiccia, University of Delaware, University of Delaware, Depatment of Food and Resource Economics, Newark, DE 19716, *lariccia@udel.edu*; Paul P.P. Eggermont, University of Delaware

Strong approximations by simple Gaussian Processes are established for the smoothing spline estimator (minus it's expectation). The result is shown to hold for splines of arbitrary order, and for both deterministic and data driven selections of the smoothing parameter. The result provides a simple method for determining the asymptotic distribution of test statistics and pivot variables, based on the spline estimator. This is illustrated by determining asymptotic confidence bands for the mean function. Also, a method of bias correction and or undersmoothing is presented.

Sharp Nonparametric Estimation and Bioapplications

✱Sam Efromovich, University of Texas at Dallas, 4525 Saint James Dr, Plano, TX 75024, *efrom@utdallas.edu*

Key Words: density, characteristic function, fMRI, microarray

Sharp nonparametric estimation is a topic devoted to finding best rates and constants for estimation of nonparametric curves under minimax approach. New theoretical results for the classical setting of the probability density and characteristic function estimation are presented. The results include lower bounds, upper bounds and oracle inequalities. Second part of the talk is devoted to some recent applications of nonparametric curve estimation methods, including wavelet estimates, in bioapplications.

Semi-Supervised Wavelet Thresholding and Applications

Kichun S. Lee, Georgia Institute of Technology; ✱Brani Vidakovic, Georgia Institute of Technology, Atlanta, GA 30332, *brani@bme.gatech.edu*

Key Words: semi-supervised, wavelet regression, denoising

Under the general regression setup, $y_i = f_i + \epsilon_i$, we are interested in estimating a possibly multivariate regression function f. The wavelet thresholding is a simple operation in the wavelet domain that selects the subset of wavelet coefficients corresponding to an estimator of f when back-transformed. We propose selection of the subset in a semisupervised fashion. A distance function appropriate for wavelet domains is proposed. The “energetic” coefficients are not independent and fall on low dimensional manifolds in the space of all wavelet coefficients. The decision to include a coefficient in the model depends not only on its magnitude but also on the distance from the coefficients labeled as “energetic” and “non-energetic.” Method’s theoretical properties are discussed and its performance demonstrated on simulations and real-life examples.

132 Importance of Predictors in Statistical Models ●

Section on Statistics and Marketing, Section on Quality and Productivity

Monday, August 4, 10:30 a.m.–12:20 p.m.

Estimation of Predictor Shares via Multinomial Parameterization: Application to Regression Coefficients and to PCA and SVD Loadings

✱Stan Lipovetsky, GfK Custom Research North America, 8401 Golden Valley Rd, Minneapolis, MN 55427, *stan.lipovetsky@gfk.com*

Key Words: regression, predictors importance, multinomial parameterization, PCA/SVD

Multiple linear aggregates with special properties of their coefficients parameterized by exponent, logit, and multinomial functions are considered. To obtain always positive coefficients of multiple regression, or coefficients with the signs of pair correlations, the exponential parameterization is applied. To attain the coefficients belonging to an assigned range the logistic parameterization is used. The coefficients equal the shares of influence themselves can be obtained using the multinomial parameterization. The work also considers multinomial logit parameterization of the eigenvectors’ elements that always yields nonnegative loadings of shares for variable aggregation in principal component analysis and in the singular value decomposition.

Maximum Entropy Choice Probabilities vs. Random Utility Choice Probabilities

✱Robert Bordley, General Motors Research Labs, 30500 Mound Road, Warren, MI 48090-9055, *robert.bordley@gm.com*; Ehsan Soofi, University of Wisconsin

Key Words: maximum entropy, discrete choice, random utility, ideal point, approximations

The widely used random utility model of discrete choice suffers from making the widely rejected assumption of customers being rational. Maximum entropy models do not make this assumption. Both maximum entropy and random utility theory lead to the logit choice model in the simplest case. But they differ in how they generalize that logit model. Since the logit model typically must be generalized to be realistic, this paper derives the generalization of the logit model suggested by maximum entropy theory. It has two advantages over the random utility models: (1) It’s interpretability as an ideal-point (or anti-ideal-point model); (2) It’s ease of estimation; (3) It’s ability to closely approximate the most popular random utility model other than the logit (i.e., the mixed logit model).

Information Importance of Predictors: Concept, Measures, Bayesian Inference, and Applications

✱Ehsan Soofi, University of Wisconsin, Business Administration Building S479, Madison, WI 53201, *esoofi@uwm.edu*; Joseph Retzer, Maritz Research; Refik Soyer, The George Washington University

Key Words: Contingency table, Exponential family regression, Invariance, Mutual Information, Linear regression, Logit

Importance of predictors is characterized by the extent to which their use reduces uncertainty about predicting the response variable, namely their information importance. Uncertainty associated with a probability distribution is a concave function of the density such that its global maximum is a uniform distribution reflecting the most difficult prediction situation. Shannon entropy is used to operationalize the concept. For nonstochastic predictors, maximum entropy characterization of probability distributions provides measures of information importance. For stochastic predictors, the expected entropy difference gives measures of information importance. Applications to various data types lead to familiar statistical quantities for various models, yet with the unified interpretation of uncertainty reduction. Bayesian inference and examples will be presented.

Importance of a Predictor When the Predictor Is Also the Dependent Variable of Interest

✱Alaattin Erkanli, Duke University, 905 West Main Street Suite 22, Brightleaf Sqaure, Durham, NC 27701, *aerkanli@psych.duhs.duke.edu*

Key Words: Markov Process, Markov Chain Monte Carlo, Latent Transition/Hidden Markov Models, Bayesian Inference

We discuss a class of transition models where the predictors are the past outcomes. Both latent (hidden) or observable variables are considered. We specifically developed Bayesian methods to estimate the order of transitions when the transitions are governed by a generalized Markov/Semi-Markov process and the number of states vary over time, using variable selection algorithms. Both simulated and real world examples are provided to illustrate these models.

133 Teaching Through Service Learning: Getting Statistics Out of the Classroom While Enhancing Learning ●▲

Section on Statistical Education

Monday, August 4, 10:30 a.m.–12:20 p.m.

Teaching Through Service Learning (SL): Getting Statistics out of the Classroom While Enhancing Learning

✱Amy L. Phelps, Duquesne University, 600 Forbes Ave, 830 Rockwell Hall, Pittsburgh, PA 15282, *phelpsa@duq.edu*; ✱Norean R. Sharpe, Babson College, 314 Babson Hall, Babson Park, MA 02457, *sharpen@babson.edu*; ✱Paul J. Roback, St. Olaf College, 206 Old Music Hall, St. Olaf College, Northfield, MN 55057, *roback@stolaf.edu*; ✱Beth Walters, Loyola College in Maryland, Department of Mathematical Sciences, 4501 N. Charles St, Baltimore, MD 21210, *ewalters@loyola.edu*

Key Words: Service Learning, Consulting Projects, Classroom Pedagogy

“Service learning refers to a method under which students apply particular concepts to real-life situations” (Furco 2003). It is a classroom pedagogy that integrates community service and reflective thinking to enhance student learning and social responsibility while providing a needed service to a community partner. The purpose of this session is to present examples of SL consulting projects and pedagogy which can be used in an introductory nonmajor or a statistics major course. Topics include finding projects, student meetings and work with projects, classroom requirements, instructor involvement, project results, and student and project feedback. The use of student reflections will be introduced to show how this aspect of SL enhances learning while giving the student a sense of community involvement and personal development. Ideas will be shared in a breakout session format.

134 Inference Issues in Biometric Data

Biometrics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

A New Confidence Interval for a Simple Effect in a 2-by-2 Factorial Experiment Utilizing Uncertain Prior Information That the Two-Factor Interaction Is Zero

✱Paul Kabaila, La Trobe University, Department of Mathematics and Statistics,, La Trobe University, Victoria 3086, Australia, Melbourne, International 3086 Australia, *P.Kabaila@latrobe.edu.au*; Khageswor Giri, La Trobe University

Key Words: confidence interval, prior information, 2-by-2 factorial experiment, frequentist, two-factor interaction

Consider a 2-by-2 factorial experiment with more than one replicate. Suppose that the parameter of interest theta is a specified simple effect and that we have uncertain prior information that the two-factor interaction is zero. We describe a new frequentist 1-alpha confidence interval for theta that utilizes this prior information. This interval has expected length that (a) is relatively small when the two-factor interaction is zero and (b) has maximum value that is not too large. This interval also has the following desirable properties. Its endpoints are continuous functions of the data and it coincides with the standard 1-alpha confidence interval when the data strongly contradicts the prior information.

Testing for Group Effect in a 2 x k Heteroscedastic ANOVA Model

✱James F. Troendle, National Institute of Child Health and Human Development, NIH, Bld 6100, Room 7B05, Bethesda, MD 20892, *jt3t@nih.gov*

Key Words: empirical likelihood, likelihood ratio test, power, simulation

An approximate empirical likelihood ratio test (AELRT) is derived for the test of group main effect in a 2 x k heteroscedastic ANOVA model. To approximate the distribution of the test statistic under the null hypothesis, simulation from the approximate empirical maximum likelihood estimate (AEMLE) restricted by the null hypothesis is used. The homoscedastic ANOVA F-test and a Box-type approximation to the distribution of the heteroscedastic ANOVA F-test are compared to the AELRT in level and power. The AELRT procedure is shown by simulation to have appropriate type I error control (although possibly conservative) when the distribution of the test statistics are approximated by simulation from the constrained AEMLE. The methodology is motivated and illustrated by an analysis of folate levels in the blood among two alcohol intake groups while accounting for gender.

Tolerance Limits for a Ratio of Two Normal Random Variables

✱Lanju Zhang, MedImmune, Inc., One MedImmune Way, Gaithersburg, MD 20878, *zhangla@medimmune.com*; Thomas Mathew, University of Maryland Baltimore County; Harry Yang, MedImmune, Inc.; K. Krishnamoorthy, University of Louisiana at Lafayette; Iksung Cho, MedImmune, Inc.

Key Words: Bioassay;, Generalized confidence interval, upper tolerance limit

The problem of deriving an upper tolerance limit for a ratio of normal random variables is addressed, when the random variables follow a bivariate normal distribution, or when they are independent. The derivation uses the fact that an upper tolerance limit for a random variable can be derived from a lower confidence limit for the cdf of the random variable. The concept of a generalized confidence interval is used to derive the required lower confidence limit. In the bivariate normal case, a suitable representation of the cdf of the ratio of the marginal normal random variables is also used, coupled with the generalized confidence interval idea. In addition, a simplified derivation is presented in the situation when one of the random variables has a small coefficient of variation. The problem is motivated by an application from bioassay.

Longitudinal Assessment of Hearing Toxicity in a Phase III Clinical Trial for Reduction of Sporadic Adenomas

✱Christine E. McLaren, University of California, Irvine, Department of Epidemiology, 224 Irvine Hall, Irvine, CA 92697-7550, *cmclaren@uci.edu*; Wen-Pin Chen, Chao Family Comprehensive Cancer Center; Sharon Fujikawa-Brooks, Chao Family Comprehensive Cancer Center; Frank L. Meyskens, Chao Family Comprehensive Cancer Center

Key Words: Clinikcal Trials, Repeated measures, GEE, Toxicity, Data and Safety Monitoring

A DSMB recommended early termination of a Phase III trial in favor of efficacy and requested methodology to assess hearing toxicity. By initial criteria, there were 25 of 136 (18.4%) in the treatment group and 12 of 123 (9.8%) taking placebo who experienced hearing loss (RR, 1.88; 95% CI, 0.99 to 3.59). Treatment groups were similar with regard to time between randomization and the outcome audiogram. Lowess-smoothed profiles showed comparable hearing changes over time. GEE methodology was applied with subjects as clusters and pure tone threshold measured at the end of the therapy as the outcome variable, showing an average of < 2 dB difference in thresholds for the two groups, below the limits of testing accuracy. For assessment of longitudinal toxicity in clinical trials, appropriate methods are necessary to estimate and compare the degree of difference between groups.

Statistical Methods for Active Extension Trials

✱Zonghui Hu, National Institutes of Health, 6700 A Rockledge Dr, Bethesda, MD 20817, *huzo@niaid.nih.gov*; Dean Follmann, National Institutes of Health

Key Words: ANCOVA, change score, pretest-posttest design

We develop analysis method for active extension trials. Under this design, patients are randomized to treatment or placebo for a period of time (period 1), then all patients receive treatment for an additional period of time (period 2). We assume a continuous outcome is measured at baseline and at the end of the two consecutive periods. If only period 1 data is available, the treatment effect estimators include the change score, ANCOVA, and maximum likelihood. We show how to extend these estimators by incorporating period 2 data. Under the assumption that the mean responses for treatment and placebo arms are the same at the end of period 2, the new estimators are unbiased and more efficient. Otherwise, the period 2 estimators may be biased. We propose the period 2 estimator as an efficient supplement but not a replacement of the classic estimators.

A Transformation Binormal Model To Estimate the ROC Curve of a Continuous Biomarker

✱Danping Liu, University of Washington, F-600 Health Sciences Building, 1705 NE Pacific Street, Seattle, WA 98195-7232, *dpliu@u.washington.edu*; Andrew Zhou, University of Washington

Key Words: ROC curve, Binormal model, Local polynomial, Transformation model

Binormal models are widely used to estimate the ROC curve of a continuous biomarker, assuming the marker values for cases and controls both follow normal distributions with different location and scale parameters. We relax this assumption and set up a transformation binormal model (i.e., the normality of the marker values is achieved by an unknown monotone transformation). A local polynomial framework is adopted to estimate the transformation function. The individual covariates could be adjusted by considering some profile kernel estimating equations. Simulation studies are conducted to examine the performance of our estimate. The proposed method is applied to a real data set arising from the Alzheimer's disease research.

Testing for Heterogeneity Among the Components of a Binary Composite Outcome

✱Janice Pogue, McMaster University, PHRI- HGH, 237 Barton Street East, Hamilton, ON L8L 2X2 Canada, *janice@phri.ca*; Lehana Thabane, McMaster University

Key Words: composite outcomes, heterogeneity, clinical trials, correlated binary outcomes

Clinical trials use composite outcomes to overcome many statistical issues, including power and inflation of type I error. However, there are problems with the interpretation of composite outcomes, in terms of possible heterogeneity of treatment effect on the components of these composites. If significant heterogeneity is found the composite outcome may be invalidated/inappropriate for use. However, tests for heterogeneity are known to lack power. The purpose of this paper is to compare the power of different tests to detect heterogeneity of treatment effect across components of composite binary outcomes. We conducted a simulation for a typical trial in cardiovascular disease to test the power of 5 different models detect treatment heterogeneity among components of a composite. Power is shown for variations in heterogeneity degree, component numbers/balance, correlation, and sample size.

135 Rethinking Established Practice ▲

Social Statistics Section, Section on Government Statistics

Monday, August 4, 10:30 a.m.–12:20 p.m.

Effects of Dependence on Meta-Analytic Tests

✱Kevin Henning, Texas Tech University, Rawls College of Business, Rawls College of Business, Area of ISQS, Lubbock, TX 79409-2101, *kevin.henning@ttu.edu*

Key Words: Meta-analysis, Hypothesis testing, dependence structures, Monte Carlo, Type I errors

Researchers frequently are interested in "analyzing the analyses" conducted by others in a particular field. Known as meta-analysis, the results of such investigations often lend credence to a particular theory or important finding. Several meta-analytic tests have been proposed in the literature, most notably the Liptak, Fisher, Sidak, and Simes combination tests. A important assumption underlying these tests is that p-values, when the result of properly conducted hypothesis tests, are i.i.d. uniform(0,1). In this paper I examine through Monte Carlo simulation the effects on the nominal Type I error rates of these tests under two types of dependence structures that could easily arise in practice: compound symmetry and AR(1) dependencies. I provide relevant discussion and suggestions for future research.

Classification Rules for Repeated Measures Data from Multiple Sources

✱Anuradha Roy, The University of Texas at San Antonio, Department of Management Science and Statistics, One UTSA Circle, San Antonio, TX 78249, *aroy@utsa.edu*; Ricardo Leiva, Universidad Nacional de Cuyo

Key Words: classification rules, jointly equicorrelated covariance structure, maximum likelihood estimates, separable means

Under the assumption of multivariate normality we study new classification rules for multiple m-variate observations over u-sites and over v-time points. We assume that the m-variate observations have "jointly equicorrelated covariance" structure and a separable mean vector. The new classification rules are very efficient in discriminating individuals when the number of observations is very small, and thus unable to estimate the unknown variance-covariance matrix. We present computation algorithms for maximum likelihood estimates of the unknown population parameters. We demonstrate the classification rules on a real data set. Our result shows that our new classification rules are far better than the traditional classification rules for small to moderate sample sizes. These classification rules have applications in biomedical, medical, pharmaceutical and many other research areas.

New Modified Chi-Squared Goodness-of-Fit Tests and Their Application in Particle Physics and Political Science

✱Vassilly Voinov, KIMEP, 2 Abai Av., Almaty, 050010 Kazakhstan, *voinovv@kimep.kz*

Key Words: Poisson distribution, Binomial distribution, Feller's distribution, modified chi-squared tests, US Supreme Court, Rutherford's experiment

Modified chi-squared tests for Poisson, binomial, and approximated Feller's distributions are discussed. A reanalysis of the classical Rutherford's experimental data on alpha decay is done. Previous analyses of the data were not correct from the point of view of the theory of statistical testing. Suggested tests show that the data contradict to both Poisson and binomial distribution and do not contradict to a precise "binomial" approximation of Feller's distribution, which accurately takes into account a counter's dead time. This gives a plausible statistically correct confirmation of the well-established exponential law of radioactive decay. An application of the suggested tests for the analysis of the number of US Supreme Court justices' appointments from 1789 up to date shows that the binomial process has the same right to exist as the Poisson one under use.

Statistical and Physical Models in Social Sciences: A New Reconciliation

✱Igor Mandel, Advanced Marketing Models, 3-58 26th street, Fair Lawn, NJ 07410, *igor.mandel@gmail.com*

Key Words: statistics, sociophysics, econophysics, mediaphysics, modeling, social sciences

More and more physicists apply different types of models borrowed from statistical physics or specifically designed for concrete purposes in social sphere. A history of relations between physics, statistics and social sciences is quite long, although it is not fully comprehended yet. It gives a spectacular picture, when basic physical and statistical concepts, that seem very remote now, were considered at some point of time in much more integrated way. However, in the last decade there are more and more evidences of the new reconciliation of two sciences. This paper outlines the main features of this process. It shows the difference between physical and statistical models and shortly describes the approach which could bridge these two paradigms—mediaphysics.

Who (Really) Are the First Boomers?

✱Howard Hogan, U.S. Census Bureau, 695 Woodland Way, Owings, MD 20736, *hhogan@census.gov*; Debbie Perez, U.S. Census Bureau; William R. Bell, U.S. Census Bureau

Key Words: births, time series

The press made a big deal when the "first person" born in 1946 applied for social security. On the other hand, many "generational analysis" guru's call everyone born in the 1940s the Boomer Generation. Since World War II did not end until August 15, 1945, one might think that the "post war" baby boom could not start much before nine months later, or mid May 1945. This paper looks at the monthly time series of births to determine the start of the Baby Boom. It uses modern series techniques to determine when there was a "break in series" that would constitute the beginning of a boom. It also will examine other data to put the beginning of the boom in context.

A Trigger for Further Auditing After a Post-Election Audit for Winner Verification

✱Vittorio Addona, Macalester College, 1600 Grand Avenue, Saint Paul, MN 55105, *addona@macalester.edu*; Katherine Lim, Macalester College

Key Words: post-election audit, winner verification, power, trigger

Post-election audits to verify electronic vote counts have received growing attention. Previous research focused on shortcomings of fixed, or tiered, percentage audits and on the development of a method to determine the audit sample size based on power. A power-based method ensures a high probability of detecting a miscounted precinct if enough exist to have altered the election outcome. If an audit reveals no miscounted precincts it is complete. The procedure to follow, however, when miscounts are found remains unclear. The detection of miscounts may not require further auditing to verify the winner. We propose a trigger for additional auditing that is based on a modified Hoeffding bound. This trigger is compared, through simulations, to one based on a bootstrap confidence interval estimate of the net gain in votes that the reported loser would obtain should a 100% audit be conducted.

A Study on Tests of Symmetry with Ordered Alternatives in Higher-Dimensional Contingency Tables

✱Ping Ye, Southern Illinois University, 1961 Evergreen Terrace Dr. w Apt 5, Carbondale, IL 62901, *pingye@siu.edu*; Bhaskar Bhattacharya, Southern Illinois University Carbondale

For a two-dimensional contingency table of probabilities, the concept of symmetry around the main diagonal is well defined. Statistical hypothesis test of symmetry versus positive bias have also been explored. For tables of higher (three or more) dimensions, however, different concepts of symmetry are available. In this study, we consider statistical inference procedures of symmetry in partial tables versus various biases in three-dimensional tables. We find the maximum likelihood estimates of the cell probabilities and the asymptotic distribution of the likelihood ratio test statistic in each case. Simulation studies are used to investigate the sizes and powers of the tests. The methodologies developed are applied on real data sets.

136 Genetic Association Studies ●

Section on Statistics in Epidemiology, Biopharmaceutical Section, Biometrics Section
Monday, August 4, 10:30 a.m.–12:20 p.m.

Incorporating Heterogeneity of Maternal Effects for Precisely Detecting Parent-of-Origin Effect

✱Jingyuan Yang, The Ohio State University, 404 Cockins Hall, 1958 Neil Ave., Columbus, OH 43210, *yj@stat.osu.edu*

Key Words: parent-of-origin effect, maternal effect, heterogeneity, genomic imprinting, microchimerism, association

Once association is detected between a marker and a disease locus, researchers may further investigate the existence of maternal and parent-of-origin (POO) effects. Maternal effect is conventionally modeled as a fixed effect, assuming that maternal variant allele(s) has (have) identical effect on any offspring. The presence within a person of maternal substances, passed from his or her mother during pregnancy, is referred to as maternal microchimerism (MMc). Recent studies reported that levels of MMc vary dramatically from person to person, which could result in heterogeneity of maternal effects. In this report, several methods of detecting POO effect are compared under different scenarios of maternal effects, and a generalized linear model is proposed to incorporate heterogeneous maternal effects in detecting POO effect. The proposed model is evaluated using simulated data.

Information Geometry, Gene-Gene, Gene-Environment Interaction, and Pathway Association

✱Momiao Xiong, The University of Texas School of Public Health, 1200 Herman Pressler, Houston, TX 77030, *Momiao.Xiong@uth.tmc.edu*; Li Luo, The University of Texas School of Public Health; Gang Peng, Fudan University

Key Words: Gene-gene interaction, Gene-environment interaction, Information theory, Genetic networks, pathway, genome-wide association studies

Despite current enthusiasm for investigation of gene-gene, and gene-environment interactions, and pathway association, the essential issue of how to detect gene-environment interaction and identify pathway association remains unresolved. In this report, we use concept in information geometry to measure and develop novel tests for gene-environment interaction and pathway association. We validate the null distribution and calculate type 1 error rates of developed statistics using extensive simulation studies. We found that the new test statistics were much more powerful than other traditional statistics under several disease models. Finally, the developed statistics were applied to a number of real examples. The results showed that P-values of the information geometry-based statistics were much smaller than that obtained by other approaches.

Functional Genetic Models for Unraveling Path from Genomic Information to Complex Phenotypes

✱Li Luo, The University of Texas School of Public Health, 1200 Herman Pressler, Human Genetics Center, School of Public Health, Houston, TX 77030, *Li.Luo@uth.tmc.edu*; John Reveille, The University of Texas Medical School; Momiao Xiong, The University of Texas School of Public Health

Key Words: functional genetic model, functional data analysis, quantitative trait, genome-wide association studies, genetic epidemiology, complex diseases

Traditional quantitative genetics has primarily studied traits as isolated and static variables. Less attention is paid to dynamic behaviors of the traits. However, in real biologic world, many quantities often change over time. Great conceptual and statistical challenges are how to characterize and investigate not only steady-state, but also dynamic behaviors of the biological processes. To meet these challenges, we propose to develop functional genetic models which take traits as random functions and to address the issues about how to develop functional genetic models that use longitudinal intermediate traits as response and predictor variables? The proposed functional genetic models were applied to genome-wide association studies of ankylosing spondylitis (AS) which successfully leads to identification of SNPs, genes and biological networks associated with AS.

Bayesian Mixture Models for Case-Control, Genome-Wide Association Studies

✱Lin Li, Cornell University, Ithaca, NY 14850, *ll323@cornell.edu*; Andrew G. Clark, Cornell University; Carlos D. Bustamante, Cornell University

Key Words: genome-wide association study, QTL mapping, dimension reduction, case-control study

In this paper, we develop Bayesian mixture models for dimension reduction in case-control genome-wide association studies. The method proposes incorporating the prior information that most genotypic markers have a negligible probability of being quantitative trait loci (QTL). Both the logistic model and the probit model are considered for multiple QTL effects on disease susceptibility, and Markov chain Monte Carlo (MCMC) techniques are developed for inference. A novel method called Score Averaging Method (SAM) is also proposed for accelerating dimension reduction. The approaches are applied to simulated data of case-control association studies from real-life genotype data, and our results suggest that the method is nearly always more powerful than single marker analyses in identifying multiple QTL.

A Multiple Testing Correction Method for Genetic Association Studies Using Correlated SNPs

✱Xiaoyi Gao, Miami Institute for Human Genomics, 1120 NW 14th Street, Miami, FL 33136, *xgao@med.miami.edu*; Joshua Starmer, The University of North Carolina at Chapel Hill; Eden R. Martin, Miami Institute for Human Genomics

Key Words: single nucleotide polymorphism, composite linkage disequilibrium, multiple testing correction, eigenvalues, principal component analysis

Correcting for multiple testing is a challenging and critical issue in genetic association studies using large numbers of single nucleotide polymorphisms (SNPs), many of which exhibit linkage disequilibrium (LD). The Bonferroni method is well known to be conservative in the presence of LD. Permutation-based corrections can correctly account for LD among SNPs, but are computationally intensive. In this work, we propose a new multiple testing correction method for association studies using principal components of SNP markers. We show that it is simple, fast and is comparable to permutation-based corrections using both simulated and real data. We also demonstrate how it can be used in whole-genome association studies to control type I error. The efficiency and accuracy of the proposed method make it an attractive choice for multiple testing of SNPs in LD with each other.

A Novel Nonparametric Test for Admixture Mapping in Unrelated Individuals Based on Functional Principal Component Analysis

✱Kun Chen, University of California, Davis, One Shields Avenue, Department of Statistics, Davis, CA 95616, *kunchen@ucdavis.edu*; Jane-Ling Wang, University of California, Davis

Key Words: Admixture mapping, Ancestry Proportion, Functional Principal Component Analysis

Admixture mapping is a method that aims to detect the location of a disease which has different risks across the ancestry populations that are mixed recently. Existing methods that test for ancestry association with the disease phenotype are either model based (McKeigue,2004) or bootstrapping a summary statistic (Pritchard, 2004). We propose a flexible model, yet preserving most information in the distribution. We model the ancestry proportion density function for each marker locus by Functional Principal Component Analysis (FPCA) model, and test the deviation of each density function from the null distribution, estimated by combining information from genome-wide loci. Simulations show that the proposed method reduces the false positive rate while attaining desirable power compared to the existing methods.

Ancestral Haplotype Inference Based on Conditional Randoms Fields and Its Application to Disease Association Test

✱Xiaolin Yin, Case Western Reserve University, Electrical Engineering and Computer Science Department, 10900 Euclid Ave, Cleveland, OH 44106, *xiao.yin@case.edu*

Key Words: Conditional Random Fields, Ancestral Haplotype, Disease Association Test

We propose a statistical model for inferring the ancestral haplotype of unrelated individuals. Our model is based on the framework of Conditional Random Fields for the observed genotype data. We propose a method to incorporate the genotype data in defining the transition feature and also investigate the parameter estimation for our proposed model. We conduct the disease association study based on the inferred ancestral haplotype.

137 Spatial Statistics: Theory

Section on Statistics and the Environment, IMS
Monday, August 4, 10:30 a.m.–12:20 p.m.

Large-Scale Structure in Multiparameter Markov Random Field Models

✱Jonathan Hobbs, Iowa State University, 2804 Stange Rd #8, Ames, IA 50010, *jonhobbs@iastate.edu*

Key Words: Markov Random Fields, Spatial Statistics, Meteorology

Markov random field (MRF) models continue to see extensive use in spatial and spatio-temporal settings. Recent developments include parameterizations for non-Gaussian natural exponential family distributions that incorporate large-scale or covariate information in a fashion that generally maintains the interpretability of parameters and marginal expectations of models with no dependence. We extend this decomposition approach to MRF models to the multiparameter exponential family case and illustrate with parameterizations for models with beta conditionals. A particular challenge is devising a parameterization that ensures the desired local dependence, but the multiparameter case offers some flexibility in other aspects. We demonstrate general strategies for estimation and compare the approach to some competing models using a meteorological data set.

On the Codispersion Coefficient for Spatial Sequences

✱Ronny O. Vallejos, Universidad de Valparaíso, Avenida Gran Bretaña 1091, Playa Ancha, Valparaiso, International V Chile, *ronny.vallejos@deuv.cl*

Key Words: Codispersion coefficient, Spatial autoregressive models, correlation coefficient

In this talk some properties of the sample codispersion coefficient for spatial and temporal sequences are discussed. The consistency and asymptotic normality are established for any error distribution admitting finite fourth moment. The performance of the codispersion coefficient is examined by Monte Carlo simulation. Approximate confidence intervals and a test of independence are addressed. To illustrate the applicability of the coefficient, a real data example is presented in the context of flammability properties of polymers.

Estimation with Extended CAR Models for Large Lattices

✱Ernst Linder, University of New Hampshire, Dept. of Math. & Statistics, Kingsbury Hall, Durham, NH 03824, *elinder@unh.edu*

Key Words: Gaussian random fields, spatio-temporal analysis, large data, approximate inference, regional climate models

We propose an extension of the usual CAR model for spatial lattice data. The extension incorporates a second spatial parameter that governs smoothness of the underlying spatial field. This model is defined via the spectral decomposition of the precision matrix of the data. For regular rectangular lattices, assuming a circulant structure, the spectral decomposition is the Fast Fourier transform (FFT) and thus provides a computationally feasible method for very large data. We consider strategies to applying this model for a variety of space - time regression and functional ANOVA type models. We provide an implementation within a hierarchical Bayesian estimation framework applied to regional climate model output data over a large grid covering North America.

Space-Time Models: A Spatially Dependent Filtering Approach

✱Peter F. Craigmile, The Ohio State University, Department of Statistics, 404 Cockins Hall, 1958 Neil Avenue, Columbus, OH 43210, *pfc@stat.osu.edu*

Key Words: geostatistical processes, discrete-time time series, nonstationarity, prediction

Many phenomena in nature are measured across space and through time. Investigating possible space-time interactions, in the presence of uncertainty, is key to understanding the science. In this talk we propose models that are defined as spatially dependent filterings of space-time innovation processes that are realizations of geostatistical processes. These models include the usual class of linear time series models, as well as standard geostatistical models, both Gaussian and non-Gaussian. The processes do not need to be stationary in time or in space, and build on the growing literature concerning nonstationary models. Leveraging both the innovations, as well as the filtering operations used to define these space-time processes, we use ideas from spatial statistics and time series to develop methods for inference and prediction in a space-time context.

Thinning and Augmenting of Point Process Patterns for Model Assessment Using Ripley's K-Function

✱Alejandro Veen, IBM T. J. Watson Research Center, 1101 Kitchawan Road, Route 134, Yorktown Heights, NY 10598, *av7000@gmail.com*

Key Words: point process models, K-function, thinning, geographic point pattern

This work discusses how thinning and augmenting point patterns can be used to asses the goodness-of-fit of point process models. After presenting the distributional properties of Ripley's K-function and some of its variants, thinning and augmenting will be applied to point patterns representing the geographic locations of villages in order to select a model that best fits the data.

A Penalized Likelihood Approach to Modeling Spatial Counts and Mapping Indices of Aggregation

✱Nels Grevstad, Metropolitan State College of Denver, Dept of Mathematics, denver, CO 80217-3362, *ngrevsta@mscd.edu*

Key Words: penalized likelihood regression, spatial aggregation, negative binomial distribution, ecological data

A model for counts data on a spatial domain is presented. The counts are assumed to follow the negative binomial distribution, with both the mean and dispersion parameter allowed to vary spatially. The mean is estimated via semiparametric penalized likelihood regression. The dispersion parameter is assumed to follow a parametric model whose parameters are estimated through minimization of the generalized approximate cross validation score of Xiang and Wahba (1996), which is simultaneously used to estimate the smoothing parameter associated with estimation of the mean. The increased flexibility permitted by allowing the dispersion parameter to vary spatially is illustrated through spatial maps of certain ecological indices of aggregation that are based on the mean and variance, such as Morisita's index of dispersion and Lloyd's index of patchiness, using real and simulated data.

The Generalized Linear Model with Spatial Variation and Multiple Covariates

✱J. Aleong, University of Vermont, Dept of Mathematics and Statistics, 24 Hills, Burlington, VT 05405, *jaleong@uvm.edu*

Key Words: Hierarchical generalized linear model, Spatial statistics, Longitudinal Study, Correlation, Model Checking.

The generalized linear models (GLM) and quasi-likelihood give a flexible framework for analyzing data generated from an exponential family of distributions which includes non-normal data with many types of correlation structures. This theory includes the analysis of discrete and categorical spatial data. Treatment effects in a designed experiment involving continuous and discrete spatial responses with covariates can be estimated and tested. Universal kriging and indicator kriging are discussed as special cases. Testing model assumptions to discriminate between correlation structures will be demonstrated. Examples are given, on comparing treatments in a designed experiment in which spatial correlation is present.

138 Graphical Displays, Maps, and Active Learning

Section on Statistical Education, Section on Statistical Graphics
Monday, August 4, 10:30 a.m.–12:20 p.m.

What's Wrong with This Picture?

✱Robert Easterling, Itinerant Visiting Professor, 51 Avenida del Sol, Cedar Crest, NM 87008, *rgeaste@comcast.net*

Key Words: Introductory Statistics, Graphical Methods

Introductory statistics classes and texts frequently start with graphical methods. In teaching students why and how to produce and interpret graphical displays, care is needed to assure that the displays do not distort or omit important information. For example, a histogram display of time series data loses the sequential nature of the data. We should avoid asking students to do such plots just to practice making histograms. I present some examples that illustrate this concern and suggest some principles to teach students along with the mechanics of creating data displays.

Why Introductory Statistics Courses Should Not Be Mapless

✱Mary Richardson, Grand Valley State University, Department of Statistics, 1133 MAK, 1 Campus Drive, Allendale, MI 49401, *richamar@gvsu.edu*; Paul Stephenson, Grand Valley State University

Key Words: statistical map, geographic information systems

Many governmental agencies and communication organizations use Geographic Information Systems (GIS) to accumulate and present data. Maps are an excellent way of displaying statistical data that comes from a geographical frame. Maps are not only aesthetically appealing, but they make it easier to identify geographic patterns in the data set. However, few introductory statistical texts and courses present maps as a means for displaying data. In this talk the authors will present a number of examples of different types of statistical maps, and the authors will discuss why these statistical maps should be included in introductory statistics courses.

Graphical Exploration and Suggestions for the Selection of a Regression Model Using Real Data

Irene Schiattino, Universidad de Chile; Claudio Silva, Universidad de Chile; ✱Carmen Acuna, Bucknell University, Moore Avenue, Lewisburg, PA 17837, *cacuna@bucknell.edu*; Rosa Montaño, Universidad de Chile; Verónica Iglesias, Universidad de Chile

Key Words: Least squares estimation, weighted least squares, maximum likelihood

The wide availability of statistical software might lead us to rush into the model fitting process and to bypass the use of efficient graphical representations of the data without taking the time to analyze the coherence and the logic of the strategies employed. Both, the modeling process and the graphical representation should consider the context induced by the statistical theory and the area of application. In this work we discuss, from an educational perspective, several approaches for fitting a regression model when heteroscedasticity is present. The work emphasizes the use of regression graphics (scatterplots, residual plots, scatterplot matrix, three-dimensional plots, smooth curves, etc.), and compares different model fitting methods. Data from two Chilean research projects are used.

Visualization of Probability Distributions

Betul Atalay, Saint Joseph's University; Agnes M. Rash, Saint Joseph's University; Deborah Lurie, Saint Joseph's University; ✱Susanna Wei, Saint Joseph's University, 5600 City Avenue, Philadelphia, PA 19131-1395, *swei@sju.edu*

Key Words: data visualization, probability distribution, interactive graphics, probability model, statistical inference, mathematical statistics

We are developing a collection of interactive probability modules to be used in calculus-based courses in probability and mathematical statistics. Each of these modules focuses on a probability distribution function or statistical concept and contains interactive graphics, data sets and suggested activities. These modules can enhance student learning by allowing students to visualize the behavior of a probability distribution function when the values of its parameters are changing. We also include many real-world data sets enabling students to explore and enhance their understanding of the relationship between theory and applications. The inclusion of suggested activities allows the student to delve deeper into the underlying concepts. The software is implemented by using JAVA. It's user-friendly and menu-driven, and it requires no programming knowledge from the user.

The Importance of a Checkpoint: Designing Active Learning Labs

✱Leigh V. Slauson, Otterbein College, 2170 Hutchman Dr., Hilliard, OH 43026, *lslauson@otterbein.edu*

Key Words: Variability, Active Learning, Conceptual Change

This talk will review what I learned about designing active learning labs during my research on student's conceptual understanding of variability. I took previous research on conceptual change in introductory statistics and applied it in a course that only had access to calculator technology for data analysis. Students were asked to use their prior knowledge to make predictions, collect and analyze data to test their predictions, and then evaluate their predictions in light of their results. Conceptual knowledge was assessed by lab questions and a CAOS pretest/posttest. I found evidence that the feedback portion of labs, where students were forced to check their reasoning with the instructor, was the most important feature of the conceptual change framework and greatly influenced student learning. I have since tried to incorporate this feature into other labs.

Project 'Project': Developing a Cumulative Group Activity for a Basic Statistics Course

✱Christopher E. Barat, Villa Julie College, 1525 Greenspring Valley Rd, Stevenson, MD 21153, *f-barat@mail.vjc.edu*; Vallory A. Shearer, Villa Julie College

Key Words: group project, basic statistics, PowerPoint presentation

As part of a restructuring of Villa Julie's Basic Statistics course, which is primarily taken by first-year students in nursing, education, and the social sciences, a semester-long group project incorporating all aspects of the course material was instituted in the Fall of 2007. The project requires student groups to gather real data, analyze it using both descriptive and inferential techniques, and summarize the results in a PowerPoint presentation at the end of the semester. Students are strongly encouraged to work with data related in some way to their potential major. The format of the project has already undergone considerable revision as we attempt to make this activity a worthwhile and meaningful culmination of our students' course experience. We will discuss our progress to date and also offer some thoughts on how such a cumulative project might be meaningfully evaluated.

Teaching with Chocolate Chip Cookies

✱Herbert Lee, University of California, Santa Cruz, School of Engineering, 1156 High Street, Santa Cruz, CA 95064, *herbie@ams.ucsc.edu*

Key Words: variability, teaching statistics, active learning

Getting and retaining the attention of students in an introductory statistics course can be a challenge, and poor motivation or outright fear of mathematical concepts can hinder learning. By using an example as familiar and comforting as chocolate chip cookies, the instructor can make a variety of statistical concepts come to life for the students, greatly enhancing learning.

139 Diagnostics and Goodness-of-Fit

Section on Statistical Computing

Monday, August 4, 10:30 a.m.–12:20 p.m.

Numerator-Denominator Plots

✱William D. Heavlin, Google, Inc, 1600 Amphitheatre Parkway, Mountain View, CA 94043, *bheavlin@google.com*; Judith Koslov, Sun Microsystems, Inc.

Key Words: change points, goodness of fit, logistic regression, Mantel-Haenszel test, randomization tests, regression diagnostics

Consider estimators that are ratios of two sums, each of n terms. Examples include proportions, regression and partial regression coefficients, and Mantel-Haenszel odds ratio estimators. Such estimators share a diagnostic plot, the N-D plot, which plots the cumulative sums of the numerator terms against their respective denominator sums. The application of N-D plots depends critically on the order in which the terms are sorted and summed. This work develops the distributional properties and generalizes N-D plots to multiply-indexed sort orders. We illustrate with a diverse range of applications: empirical segmentations over a continuous scale, goodness-of-fit assessment for black box and layered models, the concurrent validity of newly proposed metrics, and the identification of rare-event risk factors.

Outlier Detection Using the Smallest Kernel Principal Components with Radial Basis Function Kernel

Yan Shen, Temple University; ✱Alan J. Izenman, Temple University, Department of Statistics, Speakman Hall (006-00), 1810 North 13th Street, Philadelphia, PA 19122-6083, *alan@temple.edu*

Key Words: Kernel Method, Principal Component Analysis, Outlier Detection

The smallest principal components have not attracted much attention in the statistics literature. This apparent lack of interest is due to the fact that, compared with the largest principal components that contain most of the total variance in the data, the smallest principal components only contain the noise of the data and, therefore, appear to contribute minimal information. This article proposes a method for outlier detection using the smallest kernel principal components in a feature space induced by radial basis function kernel. We show that the eigenvectors corresponding to the smallest kernel principal components can be viewed as that for which the residual sum of squares is minimized, and we can use those components to identify outliers with simple graphical techniques. Simulation studies show that under univariate outlier situation, the proposed method is as good as others.

Fast Diagnostics

✱Gentiane Haesbroeck, University of Liege, Grande traverse, 12, Liege, International 4000 Belgium, *g.haesbroeck@ulg.ac.be*

Key Words: Combinatorial optimization, Outlier detection, Rayleigh test of uniformity

A general problem arising in diagnostics is to identify subsamples of given size whose deletion causes maximal changes in a statistic of interest as measured by an appropriate target function. When the size of the subset is equal to one, diagnostics are usually easily derived. However, this 'delete-one' approach may suffer from the masking effect and should be completed with 'delete-m' diagnostics when necessary. However, due to the underlying combinatorial problem (choose m out on n observations), the required task becomes repetitive and highly time-consuming. The aim of the talk is to present an algorithmic way to speed up the process when the target function can be expressed in terms of probability vectors. Several applications are illustrated: outlier detection in multivariate data, detection of influential points for the Rayleigh test of uniformity.

Far Casting Cross Validation

✱Patrick Carmack, The University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas,, TX 75390-8896, *Patrick.Carmack@UTSouthwestern.edu*; Jeffrey Spence, The University of Texas Southwestern Medical Center; Qihua Lin, The University of Texas Southwestern Medical Center; William R. Schucany, Southern Methodist University; Richard Gunst, Southern Methodist University

Key Words: model selection, correlated data, extrapolation

Cross validation has long been used for choosing tuning parameters and other model selection tasks. It generally performs well provided the data are independent, or nearly so. Improvements have been suggested which address ordinary cross validations (OCV) shortcomings in correlated data. While these techniques have merit, they can still lead to poor model selection in correlated data or are not readily generalizable to high dimensional data. The proposed solution, far casting cross validation (FCCV), addresses these problems. FCCV withholds correlated neighbors during cross validation, but uses the full data set once a model is selected. The result is a technique that stresses a fitted model's ability to extrapolate rather than interpolate. This generally leads to better model selection in correlated data sets.

Judging a Model of an Experiment

✱Robert Blodgett, U.S. Food and Drug Administration, 5100 Paint Branch Parkway, College Park, MD 20740, *Robert.Blodgett@FDA.HHS.GOV*

Key Words: Goodness of fit, imporb, rarity measure, order statistics

Unexpected and unplanned events like lab errors can occur during scientific experiments of measurements. When they occur, a model ignoring them can be very inaccurate. Two diagnostic tools are compared. One is improb which equals the sum of probabilities of all outcomes as likely as or less likely than the actual outcome. The second is a rarity measure which equals the probability of the actual outcome divided by the probability of the most likely outcome.

Mixed Model Selection Criteria Based on Statistical Computation and Leave-One-Out Method

✱Junfeng Shang, Bowling Green State University, 450 Math Science Building, Bowling Green State Univeristy, Bowling Green, OH 43403, *jshang@bgsu.edu*

Key Words: Akaike information criterion, predictive divergence criterion, Kullback-Leibler discrepancy, improved Akaike information criterion

In the mixed modeling framework, statistical computation and leave-one-out method are employed to develop an improved Akaike information criterion and the predictive divergence criterion for model selection, respectively. The selection and the estimation performance of the criteria is investigated in a simulation study. The simulation results demonstrate that the predictive divergence criterion outperforms the improved Akaike information criterion in choosing an appropriate mixed model, and the improved Akaike information criterion is less biased than the predictive divergence criterion in estimating the Kullback-Leibler discrepancy between the true model and a fitted candidate model.

Extending the Range of Application for Generalized Goodness-of-Fit Statistics

✱Leah Jager, United States Naval Academy, Department of Mathematics, USNA, 572C Holloway Road, Chauvenet Hall, Annapolis, MD 21402, *jager@usna.edu*

Key Words: goodness-of-fit, multiple testing, recursive methods, finite sample calculations

Jager and Wellner (2007) introduce a generalized family of statistics for the classical goodness-of-fit problem. The asymptotic null distribution of these statistics is derived, but is not accurate for samples sizes smaller than 1,000,000. Owen (1995) developed a method for calculating exact null quantiles of the Berk-Jones statistic in samples sizes smaller than 1,000. This method can be extended to the generalized statistics, but still leaves a range of sample sizes where the statistics cannot be accurately applied. Applications of these statistics to large-scale multiple testing problems found in genetics and genomics typically involve sample sizes in this problematic range. We discuss extension of Owen's method to the generalized statistics and to sample sizes up to 10,000. These exact results are used to form approximations in the range between the exact and asymptotic methods.

140 Advances in Experimental Design ●

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Monday, August 4, 10:30 a.m.–12:20 p.m.

Optimal Two-Level Fractional Factorial Designs for Location Main Effects with Dispersion Factors

✱Chao-Ping Ting, National Cheng-Chi University, Department of Statistics, No. 64, Sec. 2, Chih-Nan Road, Wenshan District, Taipei, 11605 Taiwan, *cpting@nccu.edu.tw*; Fu-Kai Chang, National Taiwan University

Key Words: Generator, Defining relation, Dispersion effect, Location effect

In two-level fractional factorial designs, homogeneous variance is a commonly made assumption in analysis of variance. When the variance of the response variable changes when a factor changes from one level to another, we call that factor the dispersion factor. The problem of finding optimal design when dispersion factors present is relatively unexplored, however. In this article, we focus on finding optimal designs for the estimation of all location main effects when there are one or two dispersion factors present, in the class of regular unreplicated two-level fractional factorial designs of resolution III and higher. We show that by an appropriate choice of the defining contrasts, A-optimal and D-optimal designs can be identified.

Whole-Plot Exchange Algorithms for Constructing D-Optimal Multistratum Designs

✱Christopher I. Vahl, North Dakota State University, P.O. Box 5575, Waldron 201, Dept. of Statistics, North Dakota State University, Fargo, ND 58105, *chris.vahl@ndsu.edu*; George A. Milliken, Kansas State University

Key Words: multistratum design, optimal design, split-plot, exchange algorithm

Multistratum experiments contain several different sizes of experimental units. Examples include split-plot and strip-plot designs as well as randomized block designs. Here we propose to a strategy for constructing a D-optimal multistratum design by improving a randomly generated starting design through a sequence of whole-plot exchanges. This approach preserves the design structure and simplifies updates to the information matrix. Moreover, it is applicable to any multistratum design where the largest-sized experimental unit is either a whole plot or a block. Three whole-plot exchange algorithms will be described and the results of the application of these algorithms to several design problems will be discussed.

On the Weighted Optimality of Incomplete Block Designs

✱Xiaowei Wang, Virginia Polytechnic Institute and State University, 1224 University City Blvd, D47, Blacksburg, VA 24060, *xiwang3@vt.edu*

Key Words: weighted optimality, block design, efficiency balance

For incomplete block designs, optimality has been intensively studied by researchers, in terms of various criteria such as A, D, E and MV. The classic formulation of design optimality is based on the idea that optimality functions of the information matrix are invariant to treatment permutation for every permutation matrix P. This implies equal interests in all treatments. In practice, however, there are many experiments where not all treatments are equally important. For these designs, it would be better to incorporate weight factor in design optimality calculation. In this study, we justified weight E, A and MV optimality criteria, which are based on the weighted information matrix we defined. Some sufficient conditions are derived for incomplete block designs to be weighted E, A and MV-optimal for situations where weights take only two values.

A Class of Saturated Square Designs

✱Xianggui Qu, Oakland University, 2200 N Squirrel Road, Rochester, MI 48309, *qu@oakland.edu*; Theophilus Ogunyemi, Oakland University

Key Words: Binary designs, (m,s)-optimality, row-column designs, saturated designs

A new class of square designs is proposed. These designs are saturated in terms of eliminating two-way heterogeneity with an additive model. The (m,s)-criterion is used to select optimal designs. It turns out that all (m,s)-optimal designs are binary. All (m,s)-optimal designs constructed are treatment-connected. Thus, treatment contrasts are estimable in spite of the row and column effects.

A Class of Partially Replicated Two-Level Fractional Factorial Designs

✱Paul J. Lupinacci, Villanova University, Department of Mathematical Sciences, 800 Lancaster Avenue, Villanova, PA 19085, *paul.lupinacci@villanova.edu*; Joseph G. Pigeon, Villanova University

Key Words: Fractional Factorial Designs, G2-Aberration, Generalized Resolution, Uniformity, Hadamard Matrix

Resolution III fractional factorial designs permit the uncorrelated estimation of all main effects under the assumption of no interactions among the factors and are often used in the initial phases of an experimental program to screen factors. Even in these experimental situations, it may be desirable to include partial replication of the design, thereby allowing for the unbiased estimation of the experimental error while maintaining orthogonality of the main effects. A new class of partially replicated two-level fractional factorial designs based on Hadamard matrices is proposed. A method of constructing these designs is given and the relative performance of the designs is evaluated by comparing them with competing fractional factorial designs using generalized resolution, G2-aberration, and uniformity.

Nested Latin Hypercube Designs

✱Zhiguang (Peter) Qian, University of Wisconsin-Madison, Department of Statistics, 1300 University Ave, Madison, WI 53706, *peterq@stat.wisc.edu*

Key Words: design of experiments, space-filling designs, computer experiments, industrial statistics, numerical integration, computational statistics

Computer experiments with different levels of accuracy have become prevalent in practice. In this talk, I will discuss a powerful yet simple approach to the construction of a new type of design called nested Latin hypercube designs for such experiments.

Equivalence of Symmetric Factorial Experiments Having Both Qualitative and Quantitative Factors

✱Tena I. Katsaounis, The Ohio State University, Bromfield 117-3A, 1680 University Drive, Mansfield, OH 44906, *katsaounis.1@osu.edu*; Angela M. Dean, The Ohio State University

Key Words: combinatorial equivalence, geometric equivalence, symmetric factorial design, experimental design, design equivalence

The problem of equivalence of symmetric factorial designs with both qualitative and quantitative factors is considered. Two fractional factorial designs with qualitative and quantitative factors are said to be equivalent if one can be obtained from the other by reordering the runs, relabeling the factors of the same type, relabeling the levels of the qualitative factors and reversing the levels of the quantitative factors. A necessary and sufficient criterion for detecting equivalence or nonequivalence of such factorial designs is described. Some screening criteria of nonequivalence are also discussed. Results can be extended to designs with factors having different number of levels.

141 Sign and Rank Methods: New Applications and Updated Perspectives ●

Section on Nonparametric Statistics
Monday, August 4, 10:30 a.m.–12:20 p.m.

Nonparametric and Parametric Multivariate Tests

✱Arne Bathke, University of Kentucky, Department of Statistics, 875 Patterson Office Tower, Lexington, KY 40506-0027, *arne@uky.edu*; Solomon W. Harrar, University of Montana

Key Words: Multivariate Data, Multiple Endpoints, Nonparametric Tests, Rank Based Methods, High Dimension, Ordinal Data

Multivariate data appear naturally in the life sciences because usually more than one response variable is of interest (“multiple endpoints”). We look at different ways to analyze multivariate data, and at different asymptotic settings: Large n (sample size), large k (number of samples), or large p (dimension). Methodological advancements in recent years have made it possible to analyze many multivariate data sets using completely nonparametric methods. We compare different inference methods, focusing on those that are invariant under monotone transformations of the different variables, and that can be used when the data have a mix of ordinal and quantitative response variables.

Using Shrinkage in an Exact Test for Clustered Binary Data with Small Numbers of Clusters

✱Patrick Gerard, Clemson University, Dept of Applied Economics and Statistics, 291 Barre Hall, Clemson, SC 29634, *pgerard@clemson.edu*; William R. Schucany, Southern Methodist University

Key Words: sign test, bootstrap, permutation

The sign test is often employed with independent binary data. When binary data are clustered, the same hypothesis may be of interest, but the standard sign test performs poorly due to the dependence caused by clustering or blocking. Gerard and Schucany (2007) investigated a new test incorporating both permutation and an exact parametric bootstrap (EPB) to be used when the numbers of clusters are small. In this study we modify the EPB test using shrinkage of cluster level success probabilities. We compare the new test to the EPB as well as a classic permutation test for numbers of clusters ranging from 5 to 10 inclusive. We find that for most combinations of alternatives and correlations the new shrinkage test is the most powerful.

Wilcoxon-Mann-Whitney or T-Test? On Assumptions for Hypothesis Tests and Multiple Interpretations of Decision Rules

✱Michael P. Fay, National Institute of Allergies and Infectious Diseases, 6700B Rockledge Dr. MSC 7609, Bethesda, MD 20892-7609, *mfay@niaid.nih.gov*

Key Words: Hypothesis test, Wilcoxon rank sum test, Behrens-Fisher problem, Welch's t-test, permutation test, Location shift

Traditionally, a decision rule is coupled with only one set of assumptions which demarcates the null and alternative hypotheses. Here we associate a single decision rule with several sets of assumptions, each of which can define a valid or approximately valid test when associated with the decision rule. We call such a decision rule, a multiple perspective decision rule (MPDR), with each perspective providing a different interpretation to the same accept/reject decision. This allows the consumer of the decision rule to try on several sets of assumptions depending on perspective, all without recalculating the p-value associated with the decision rule. We explore this idea by reviewing the many hypotheses that have been associated with the two sample t-tests and the Wilcoxon-Mann-Whitney test, and show how the MPDR outlook may be helpful in this situation.

Recent Advances in Informative One- and Two-Sample Goodness-of-Fit Testing

✱Olivier Thas, Ghent University, Coupure Links 653, Gent, 9000 Belgium, *olivier.thas@UGent.be*; John Rayner, University of Newcastle, Australia; Donald J. Best, University of Newcastle

Key Words: hypothesis testing, goodness-of-fit, semiparametric, Neyman smooth test, nonparametric density estimation

Consider testing the one and two sample null hypothesis. A common criticism to many tests is that at the rejection of the null, they do not give an indication as to how the data deviates from the null hypothesis. With informative tests we mean tests that do tell in what direction the null hypothesis is rejected (e.g., difference in mean, variance, skewness, ...). Components of Neyman's smooth tests are well known for their informative diagnostic nature. For the one-sample problem their properties have been well studied in a semiparametric framework, but the test statistics need rescaling. In this paper we show how informative analyses can result from exploiting the relation between a data-driven smooth tests and nonparametric density estimation, and this both for the one and the two sample problems. Finally we illustrate these methods on some real data sets, using a new R package.

A Nonparametric Test of Independence Between Response and Covariate Adjusted for Treatment Effect

✱Siti Tolos, Kansas State University, Dickens 101, Kansas State University, Manhattan, KS 66506, *stolos@ksu.edu*; Haiyan Wang, Kansas State University

Key Words: Nonparametric covariate effects, Fully nonparametric model, Nearest neighborhood, Fixed window size, Measure of dependency

Effective methods for detecting general dependency among variables are in high demand, especially in high dimensional post genome data due to the complex nature of the data. Existing methods mainly focused on testing for mutual independence. In this paper, we will discuss a nonparametric test of independence between response and covariates after adjusting for nonparametric treatment effects. The test statistic is constructed using moment methods with a fixed number of nearest neighbors as pseudo replicates. The asymptotic distribution of the proposed test statistic will be presented, followed by simulation studies. Potential applications in obtaining initial connected graphs for constructing gene regulatory networks will be discussed.

Heteroscedastic, Unbalanced, Two-Fold, Nested Model When the Number of Subclasses Is Large

✱Shu-Min Liao, The Pennsylvania State University, 326 Thomas Building, Department of Statistics, University Park, PA 16802, *sxl340@psu.edu*; Michael G. Akritas, The Pennsylvania State University

Key Words: Nonparametric, Nested models, Heteroscedastic, Unbalanced design, Subclass effects, Asymptotic theory

At 2007 JSM, we have showed that in the homoscedastic balanced two-fold nested model, the classical normality-based F-test procedure for the sub-class effect is asymptotically valid. This year, we are going to present our latest findings about the two-fold nested model under heteroscedasticity and unbalanced design. The testing of the subclass effects when the number of sub-classes goes to infinity is still the core of our talk. We will show, via theoretical derivation and numerical evidence, that the classical F-test procedure is very sensitive to departures from homoscedasticity in both of balanced and unbalanced cases. As a consequence, two new test statistics (unweighted and weighted) are proposed for the heteroscedastic unbalanced cases, and their asymptotic distributions, both under the null and alternative hypotheses, are then established.

Ranking of Graph Data with Kernelized Spatial Depth

✱Cuilan Gao, University of Mississippi, Department of Mathematics, Oxford, MS 38655, *cgao@olemiss.edu*; Xin Dang, University of Mississippi

Key Words: Hibert space reproducing kernel, Laplacian matrix, Rank, Spatial depth, Spectral graph

In many areas of science, data are given in graphs, Where vertices correspond to objects and edges encode similarity among objects. Ranking graph data has been an important issue both theoretically and practically, driven by the demand of fast and efficient information processing such as the search engine in computer technology. In this talk, we propose a robust ranking method for graph data based on the recently introduced kernelized spatial depth function (Chen, Dang, Peng, Butler (2007)). KSD generalizes the spatial depth function via a Hilbert space reproducing kernel. We compare the proposed KSD-based ranking with other existing rankings. Ours demonstrates competitive performance mainly because the KSD successfully captures global and local structure of a graph.

142 Bayesian Applications in Social Sciences, Public Health, and Sports

Section on Bayesian Statistical Science, Section on Health Policy Statistics

Monday, August 4, 10:30 a.m.–12:20 p.m.

Rasch and Mixture of Generalized Linear Mixed Models for Analysis of Aphasic Deficits of Syntactic Comprehension

✱Roee Gutman, Harvard University, Science Center, Statistics Department, 1 Oxford Street, Cambridge, MA 02138-2901, *rgutman@fas.harvard.edu*; Gayle DeDe, Boston University; David Caplan, Massachusetts General Hospital; Jun S. Liu, Harvard University

Key Words: Rasch Model, Mixture models, Generalized Linear Mixed Models, Bayesian Analysis

Aphasia is a loss of the ability to produce and/or comprehend language, due to injury to brain areas responsible for these functions. Performance of aphasic patients in comprehension tests is traditionally related to personal ability and to questions' difficulty. Thus, the Rasch model should be adequate to analyze these tests. We modeled the way aphasic patients process sentence types, and their ability to accomplish tasks using Rasch model and Rasch model that incorporates task and sentence grouping effect. These models were examined using psychologically meaningful statistics and were found to be inadequate. We propose a mixture of generalized linear mixed models that cluster patients into similar response patterns and abilities. The mixture model better describes the experimental results of the performance of aphasic patients, and still incorporates the essence of the Rasch model.

Modeling Long-Run Macroeconomic Growth with Gaussian Graphical Models

✱Alex Lenkoski, University of Washington, *lenkoski@stat.washington.edu*; Adrian Dobra, University of Washington; Theo Eicher, University of Washington

Key Words: Model Selection, Graphical Models, Stochastic Search

We consider modeling determinants of long-run macroeconomic growth in a Bayesian model averaging context. Historically, there have been a variety of factors that have led to growth and the need to incorporate model uncertainty has been recognized by Sala-i-Matin (1997), Fernandez et al. (2001) and Eicher et al. (2007). We build on these results by considering growth from the viewpoint of graphical models, which introduce greater parsimony in model choice than regression models. Model search is significantly harder in this context and we review the use of the stochastic search algorithm MOSS, as developed by Dobra and Massam (2008) and Lenkoski et al. (2008). We then extend the analysis considerably by showing that the technique of instrumental variables can naturally be incorporated into graphical model search, thereby confronting endogeneity and model uncertainty simultaneously.

A Bayesian Time Series Model for Death Penalty Public Opinion

✱Kenneth E. Shirley, Columbia University, Applied Statistics Center, 2700 Broadway Apt 4N, New York, NY 10025, *ks2632@columbia.edu*; Andrew Gelman, Columbia University

Key Words: Multilevel models, Death penalty, Bayesian computing, Sample survey, Time series models

Since 1974, the General Social Survey, given annually, has asked respondents whether they support the death penalty. We model 20+ years of these individual-level binary outcomes using a multilevel logistic regression model. The model includes explanatory variables like demographic information on individuals, as well as indicators of which state the respondent lives in. We are primarily interested in modeling trends within and between states across time; to do so, we explore a variety of autoregressive models. The statistical challenges in this problem include how to model the time series component of public opinion, how to incorporate a large number of high-level interactions without overfitting, and how to perform the Bayesian computation to fit a complicated model to a large data set.

Probabilistic Projections of Fertility Using a Bayesian Hierarchical Model

✱Leontine Alkema, University of Washington, Statistics Department, *alkema@u.washington.edu*; Adrian E. Raftery, University of Washington; Patrick Gerland, United Nations Population Division; Samuel Clark, University of Washington; Francois Pelletier, United Nations Population Division

The United Nations Population Division produces estimates and projections of the total fertility rate for all countries in the world every two years. For countries with fertility above replacement level, future levels are projected based on a parametric function. We develop a Bayesian hierarchical model for producing probabilistic projections of fertility and give results for a number of Asian and African countries.

Joint Modeling of Birth Weight and Gestational Age

✱Scott L. Schwartz, Duke University, Department of Statistical Science, Box 90251, 214 Old Chemistry Building, Durham, NC 27708-0251, *scott.schwartz@stat.duke.edu*; Alan Gelfand, Duke University

Key Words: Mixture Model, Bivariate Mixture Model, Interval Censoring, Bayesian Modeling, Birthweight, Gestational Age

Customary modeling for birth weight (BW) is conditional on gestational age (GA). However, joint modeling of GA and BW offers increased flexibility and interpretation in studying relevant birth outcomes (e.g., small for GA). First, we model the bivariate distribution of BW and GA using mixtures of bivariate normals. Second, due to the interval censoring associated with GA, we introduce this mixture model as a latent specification. Finally, we introduce well-established risk factors in the form of a bivariate regression to facilitate interpretation of the joint model. The mixture modeling enables us to capture the clearly non-Gaussian distributional shapes observed in histograms. The model is applied to a portion of the 1990-2006 NC Detailed Birth Record (N=1862405) through CEHI at Duke University and is fitted using Bayesian methodology and MCMC approaches.

Improved Mixing with Simulated Annealing in Bimodal Footrace Finishing Times

✱Matthew S. Shotwell, Medical University of South Carolina, 135 Cannon Street, Suite 303, Charleston, SC 29425-8350, *shotwelm@musc.edu*; Paul J. Nietert, Medical University of South Carolina; Elizabeth Slate, Medical University of South Carolina

Key Words: Bayesian, Mixture Model, Simulated Annealing, Cooper River Bridge Run, 10k Race

The Cooper River Bridge Run is the third largest 10k foot race in the southeastern US. In 2007, there were more than 28,000 finishers. The distribution of finishing times (2007) is bimodal and varies with the participants' age and gender. The two modes are presumably due to finishing times for walkers and runners respectively. A mixture distribution is used to model these data and their covariate dependence. The typical mixture parameterization assigns a discrete latent variable to each observation which is treated as an unknown parameter. For large data sets, the size of the latent variable space restricts mixing in Gibbs sampling. A method derived from the concept of simulated annealing is used to promote early mixing of the latent variables. The effectiveness of this methodology is assessed. Finishing time statistics are summarized in light of the model.

Reweighting the Bowl Championship Series

✱Susan Buchman, Carnegie Mellon University, 5000 Forbes Avenue, Baker Hall 132, Pittsburgh, PA 15213, *sbuchman@stat.cmu.edu*; Joseph B. Kadane, Carnegie Mellon University

Key Words: football, MCMC, weighting expert opinion

The majority of statistical work on college football's Bowl Championship Series (BCS) has involved proposing or categorizing computer ratings of teams. Computer algorithms, a coaches' poll, and a media poll make up the three ratings systems that are currently equally weighted to produce an overall BCS rating, which ultimately determines which schools will compete in lucrative post-season BCS bowls. We focus on investigating the performance of the BCS as implemented for the 2004, 2005, and 2006 seasons to determine whether equal weights are appropriate. Our Bayesian analysis shows that while the posterior mode places more than half the weight on the media poll, the 95% HPD credible interval contains the equally weighted scheme. We relate our work to the ongoing controversies over the BCS.

143 Methods in Linear and Generalized Linear Models ●

ENAR, Biometrics Section
Monday, August 4, 10:30 a.m.–12:20 p.m.

Tuning Parameter Selections for Penalized Likelihood Functions

✱Yiyun Zhang, The Pennsylvania State University, 331A Thomas Building, University Park, PA 16802, *yiyun@psu.edu*; Runze Li, The Pennsylvania State University; Chih-Ling Tsai, University of California, Davis

Key Words: AIC, BIC, GIC, LASSO, Nonconcave penalized likelihood, SCAD

We apply the nonconcave penalized likelihood approach to obtain variable selections as well as shrinkage estimators. This approach relies heavily on the choice of regularization parameter, which controls the model complexity. We propose employing the generalized information criterion, encompassing the commonly used AIC and BIC, for selecting the regularization parameter. Our proposal makes a connection between the classical variable selection criteria and the regularization parameter selections for the nonconcave penalized likelihood. We show that the BIC-type selectors enable identification of the true model consistently, and the resulting estimator possesses the oracle property in the terminology of Fan and Li (2001). In contrast, the AIC-type selectors are asymptotic loss efficient under some mild conditions. Our simulation results confirm these theoretical findings.

Cumulative Logit—Poisson and Cumulative Logit—Negative Binomial Compound Regression Models for Count Data

✱Mark VanRaden, National Institutes of Health, 6700B Rockledge Dr., MSC 7609, Bethesda, MD 20892, *mvanraden@niaid.nih.gov*; John M. Lachin, The George Washington University

Key Words: count data, poisson, regression, hurdle, negative binomial, ordinal regression

Count data are often analyzed by Poisson regression, generalized to negative binomial (NB) models in overdispersed cases. When zeroes are much too frequent or infrequent, hurdle models use a binary model to predict 0 vs. >0 and a count model (e.g. conditional Poisson or NB) for the exact positive count. Model departures might still occur for very low counts. Thus we extend the binary part to a cumulative logit (CL) ordinal regression model, the conditional count model predicting exact counts within the CL's highest category, say count>L where L>0 (case L=0 is a hurdle model). Exposure time is incorporated. Parameters estimates by ML are asymptotically normal. A simple type of effect sharing between the two parts can reduce dimension. An individual Pearson type statistic crudely assesses fit. A proper, categorized chi square fit test is possible. Models were readily fit in application.

Comparison of Designs for Correlated Response Models with an Unknown Dispersion Matrix

✱Sourish Saha, GlaxoSmithKline, Collegeville, PA , *sourish.c.saha@gsk.com*; Andre Khuri, University of Florida

Key Words: correlated response models, dispersion matrix, power of multivariate tests, quantile dispersion graphs

This talk deals with the comparison of designs for correlated response models with an unknown variance-covariance matrix. The novelty of the proposed approach lies in combining the “Quantile Dispersion Graphs (QDGs)” technique with the power of the multivariate test for the equality of parameters from such models. The approach is based on considering the quantiles of a certain criterion function (namely, power of three multivariate tests) on concentric surfaces within a particular region of the so-called alternative space. The dependence of these quantiles on the unknown values of the variances and covariances obtained from the variance-covariance matrix, is depicted by plotting the so-called QDGs of the criterion function. These plots provide a clear assessment of the magnitude of the power value associated with a given design. A numerical example is presented for illustration.

Demonstration of User-Written Software in Stata, R, and Matlab for Analysis of a Familial Correlation Structure in a Quasi-Least Squares Analysis of Weight Loss in the SHARE Study

✱Xiaoying Wu, University of Pennsylvania, 501 Blockley Hall, SOM CCEB, 423 Guardian Drive, Philadelphia, PA 19104, *xwu2@mail.med.upenn.edu*; Jichun Xie, University of Pennsylvania; Sarah Ratcliffe, University of Pennsylvania; Shiriki Kumanyika, University of Pennsylvania; Justine Shults, University of Pennsylvania

Key Words: quasi-least squares, longitudinal data, correlation structure, familial structure, intervention analysis

The SHARE study was designed to facilitate weight loss in African American men and women. In this presentation I describe a quasi-least squares (QLS) analysis with a familial correlation structure to describe the pattern of association among the repeated measurements on participants in SHARE. Application of the familial structure improves our ability to understand the role that social support plays in weight loss. My is to describe implementation of our user written software in Stata, R, and Matlab for implementation of the familial correlation structure in QLS. I will describe where to obtain our software and how to apply it in analysis of longitudinal data that previously might have been analyzed using generalized estimating equations (GEE) with one of the simpler correlation structures that is available for GEE.

Variable Selection in Models with Block Structures

✱Dongyu Lin, University of Pennsylvania, 3730 Walnut Street, 400 Jon M Huntsman Hall, Philadelphia, PA 19104, *dongyu.lin@gmail.com*; Dean Foster, University of Pennsylvania

Key Words: Data Mining, Variable Selection, Block Structure, GEE, Tukey's approach

Many variable selection algorithms include variable ranking as a principle because of its scalability and efficiency in data mining. Most of the ranking methods, however, are derived from an independence assumption of the observations, which is not necessarily true in reality, and may cause some misleading results by reason of the misspecified models. But some times observations can be grouped into blocks, each block of which is independent of the other blocks. We will look into two approaches to analyze such data: the sandwich covariance matrix estimator, which works well for linear models; and a method of Tukey, which is easier and possibly more general. Here we assume the design matrix is stochastic and blocks independently follow the same distribution, under which we provide extended interpretations and asymptotical equivalence of the two approaches.

Symbolic Data Analysis: Interval-Valued Data Regression

✱Wei Xu, The University of Georgia, Dept of Statistics, 300 Rogers RD APT R302, Athens, GA 30605, *xuwei@uga.edu*

Key Words: symbolic data, interval-valued, linear regression

In statistics, usually data are formatted as single values. However, what if the data are rather represented by lists, intervals, histograms or even distributions? Symbolic data analysis was introduced to deal with these kinds of data. Among symbolic data, interval-valued data are the most commonly seen in application. Recently, different approaches have been introduced to analyze interval-valued data, including linear regression, principal component analysis and clustering, etc. My research focuses on interval-valued data regression. The study begins with the concept of symbolic data, definition of symbolic interval-valued data, its descriptive statistics and current linear regression approaches. It then seeks to introduce a symbolic covariance method. By utilizing a symbolic covariance function, outlier identification, model diagnosis and confidence interval methods are also discussed.

144 Econometric Modeling, Robustness, and Heavy Tails

Business and Economics Statistics Section, Section on Nonparametric Statistics

Monday, August 4, 10:30 a.m.–12:20 p.m.

Small Sample Performance of Instrumental Variables Probit Estimators: A Monte Carlo Investigation

✱Lee C. Adkins, Oklahoma State University, 303 Department of Economics, College of Business, Stillwater, OK 74074, *lee.adkins@okstate.edu*

Key Words: binary choice, Instrumental variables, probit

The statistical properties of several commonly used estimators of an endogenous probit model are studied using Monte Carlo simulation. Bias, standard error, and the properties of significance test are examined. Various dimensions that affect the performance of estimators of this model are examined, including sample size, proportion of observations where y=1, correlation between instruments and the endogenous variable, the correlation between the endogenous variable and the equation's error, and the relative variability of the endogenous regressor(s) and the equation's error. The estimators are used in an example that examines the effect of managerial incentives on the use of foreign-exchange derivatives by bank-holding companies.

Use of Optimal Instrument Variables for Estimating Spatial Modeling Parameters with Application to Crime Survey Data

✱Krista Collins, Statistics Canada, 100 Tunney's Pasture Drive, Ottawa, ON K1A 0T6 Canada, *krista.collins@statcan.ca*; Avinash C. Singh, Statistics Canada

Key Words: spatial autocorrelation, spatial lag model, estimating functions, instrument variables, crime data

The problem of using instrument variables in estimating regression and correlation parameters arises in the context of spatial modeling under a quasi-likelihood framework. The reason for this is that in the joint modeling of spatial lag dependence and regression, the dependent variables also appear as regressors in the model. Existing solutions to this problem do not use optimal instrument variables. The main purpose of this paper is to show how optimal instruments can be defined using estimating functions (Singh and Rao 1997) and to examine what gains in efficiency are obtained with this approach. The methodology is illustrated using simulations as well as by application to an analysis of urban neighborhood crime rates in the Canadian cities of Montreal and Edmonton.

A Robust LOESS Smooth Approach to Problem Bank Identification

Randall Kimmel, Kent State University; ✱David Booth, Kent State University, Dept. of Mngmt & IS, Kent, OH 44242, *dbooth@kent.edu*; Stephane Booth, Kent State University

Key Words: robust, loess, problem bank, smoothing, outlier

Continuing bank failures point to the need for early warning problem bank identification models to guide the actions of regulators and investors. Several models have been shown to work well, but most require extensive data preparation/manipulation and custom computer programs to analyze the data, which hinders widespread adoption. In this paper, we show that Robust LOESS (Locally Weighted Scatter Plot Smooth, a type of Local Regression Smoothing), which requires minimal data preparation and can be run in many off the shelf statistical packages such as SAS or SPSS, can be just as effective an early warning system.

Finite Sample Tests for the Composite Hypothesis of Stable Law

✱Thomas R. Willemain, Rensselaer Polytechnic Institute, 110 EIghth Street, CII-5015, Troy, NY 12180, *willet@rpi.edu*; Albert S. Paulson, Rensselaer Polytechnic Institute; Xin R. Zhong,

Key Words: stable law, finance, mixture model, characteristic function

Let a random variable be putatively from a stable law. Since sums of stable variates with index alpha are again stable with index alpha, information concerning sums of sample values is captured in powers of the empirical characteristic function (ecf). Scale invariant and nuisance free estimates of alpha are obtained via an application of the Paulson and Delehanty estimation procedure. Excessive differences between the original sample estimate of alpha and the estimates of alpha based on the ecf of sums of values imply rejection of the null hypothesis. We used extensive simulation to produce percentage points of the test for sums of 3 and 10 values and sample sizes 50, 100, 200, 400, 1000. Simulation shows that the test has remarkable power, especially against finite variance alternatives. Financial applications are provided.

Bootstrap Prediction Intervals for Large Financial Data Sets

✱William L. Seaver, The University of Tennessee, Statistics, Operations, & Mgt. Sci., College of Business Administration, Knoxville, TN 37996-0532, *wseaver@utk.edu*; Donglin Wang, The University of Tennessee

Key Words: bootsrapping, large data sets, financial data, modeling prediction intervals

In the mortgage lending industry, analysts and consultants frequently use the prediction intervals in assessing note rates or annual percent rates charged on a loan. The size of the populations can be in the millions. As a rule, prediction intervals for individual borrowers have been computed by classical multiple regression methods. However, these prediction intervals have tended to be incorrect since the residuals are not usually normal. On the other hand, to bootstrap the prediction intervals for such large populations or samples is not practical or feasible. The intent of this research is to look at the differences between the classical and the bootstrapped prediction intervals and to suggest a way to model the prediction intervals for large data sets when normality of residuals is questionable.

What Has Been the Government's Role in the Erosion of Auto Sales by U.S. Firms?

✱Ronald Tracy, Oakland University, *tracy@oakland.edu*

Key Words: CAFE standards, gas tax, econometric model, gas mileage, energy policy, auto sales

Unlike other major industrialized countries the US has used legislation (through the CAF… standards) to improve domestic fuel economy rather than a gas tax. This means that foreign automakers and US automakers respond to different market signals. When real gasoline prices hit all-time lows in the late 1990's, the US consumer wanted fast, large cars, and US auto sales boomed. However, when gasoline prices started to increase in the early 2000's US auto sales began to suffer. This study models the affect CAF… standards have had on both domestic and foreign automobile sales in the US. It shows that while the US's approach was beneficial to domestic companies when gasoline prices were low, it now disadvantages domestic auto companies. The model shows, that although the domestic companies have played a role in getting themselves in their current situation, so has US energy policy.

Competitiveness of Different Groups of Italian Enterprises Detected Using Robust Classification Methods

Luigi Biggeri, Italian National Institute of Statistics; ✱Matilde Bini, University of Florence, Viale Morgagni, 59, Department of Statistics, Florence, 50134 Italy, *bini@ds.unifi.it*; Filippo Oropallo, Italian National Institute of Statistics

Key Words: Competitiveness, efficiency, performance measures, productivity, robust cluster methods

In the recent years there as been a reduction of the competitiveness of the Italian industrial system. The main problem is the insufficient development of the total factor productivity related to the low investment in innovation of processes, products and organizational structure of the enterprises. The aim of this paper is to analyze the competitiveness for different groups or clusters of enterprises detected using adequate indicators of market and product specialization, efficiency and opening to the international markets. The clusters of enterprises are identified by a robust classification analysis. The results obtained using recent data supplied by the Italian National Institute of Statistics are presented.

145 Causal Inference ▲

Biometrics Section, Biopharmaceutical Section, Section on Health Policy Statistics, Section on Statistics in Epidemiology

Monday, August 4, 10:30 a.m.–12:20 p.m.

General Theory for Sufficient Cause Interactions

✱Tyler VanderWeele, The University of Chicago, Department of Health Studies, 5841 S. Maryland Ave., MC 2007, Chicago, IL 60637, *vanderweele@uchicago.edu*

Key Words: Causal inference, Interaction, Sufficient cause, Counterfactuals, Potential outcomes, Synergism

The sufficient-component cause framework makes reference to the actual causal mechanisms, referred to as sufficient causes, involved in bringing about an outcome. When two or more binary causes participate in the same causal mechanism, synergism is said to be present. Synergism sometimes cannot be identified from data; in cases in which data do imply that synergism is present, a sufficient cause interaction is said to be present. It is shown that any set of potential outcomes can, within the sufficient-component cause framework, be replicated by a set of sufficient causes. Empirical and counterfactual conditions are given which imply the presence of an n-way sufficient cause interaction both with and without assumptions of monotonicity. The theory and methods developed in this paper constitute empirical tests for the joint presence of two or more causes in a single causal mechanism.

Improving the Efficiency and Robustness of Doubly Robust Estimator for Estimation of Population Mean Response in the Presence of Incomplete Data

✱Weihua Cao, North Carolina State University, Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203, *wcao5@ncsu.edu*; Anastasios Tsiatis, North Carolina State University; Marie Davidian, North Carolina State University

Key Words: causal inference, doubly robust estimator, enhanced propensity score model, outcome regression, weighted least square

Considerable recent interest has focused on doubly robust (DR) estimators for a population mean response in the presence of incomplete data. DR estimators require modeling both the propensity score (PS) and the outcome regression (OR). The usual DR estimator may yield severely biased inferences if neither the PS model nor the OR model is correctly specified, and this bias is exacerbated if estimated PSs are close to zero for some observations. For arbitrary nonlinear OR models, we propose an efficient DR estimator that takes into account estimation of the PS model parameters. We also propose use of an alternative, "enhanced" PS model. We demonstrate that the resulting DR estimators yield comparable or improved performance relative to existing DR estimators, even with estimated PSs close to zero.

Semiparametric Estimation of Treatment Effects on an Outcome Measured After a Post-Randomization Event, with Missing Outcome Data

✱Yuying Jin, University of Washington, Department of Biostatistics, Seattle, WA 98109, *jinyu@u.washington.edu*; Peter B. Gilbert, Fred Hutchinson Cancer Research Center

Key Words: Causal Inference, Missing at Random, Principal Stratification, Post-treatment Selection Bias

In randomized placebo-controlled HIV vaccine trials, it is of interest to evaluate the effect of vaccination on HIV viral load or on other outcomes measured after HIV infection. Principal stratification methods can assess the average causal effect (ACE) in subjects that would be infected regardless of assigned treatment. However, the always-infected stratum is unknown. Available sensitivity analysis approaches for estimating the ACE depend on a missing completely at random mechanism, which is often untenable. To accommodate missing at random data, we integrate the previous methods with the robust likelihood-based method of Little and An (2004), which uses penalized splines of the propensity score. The new method does not require the monotonicity assumption of "no vaccine harm." It is evaluated in simulations and applied to analyze an HIV vaccine trial.

Average Causal Effect Estimation Allowing Covariate Measurement Error

✱Yi Huang, University of Maryland, Baltimore County, Dept. of Mathematics and Statistics, 1000 Hilltop Circle, Baltimore, MD 21250, *yihuang@umbc.edu*; Karen Bandeen-Roche, Johns Hopkins Bloomberg School of Public Health; Constantine E. Frangakis, Johns Hopkins University

Key Words: Causal Inference, Finite Mixture Model, Nondifferential Measurement Error, Propensity Score Subclassification, Balancing Criterion, Observational Study

In many applications of the propensity score subclassification approach, the underlying true covariates, which might confound the average causal effect (ACE) assessment of a binary treatment on health outcomes, are not directly observable, but rather are measured with error. A naive approach is to use the observed covariates to estimate propensity scores and then use these to subclassify and estimate ACEs. The work we report introduces appropriate causal assumptions for applications with errors-in-covariates and demonstrates that the naive approach for ACE estimation typically produces biased results. We also propose a flexible finite mixture framework for ACE estimation reflecting a covariate-balancing criterion in a joint likelihood, which unifies subgroup membership assessment and subgroup-specific treatment effect evaluation allowing for measurement error in the covariates.

Sensitivity Analysis of the Effect of Treatment Received in a Principal Stratification Framework

✱Corwin Zigler, University of California- Los Angeles, 628 N. Vista St., Los Angeles, CA 90036, *zigler@ucla.edu*; Tom Belin, University of California, Los Angeles

Key Words: Principal stratification, Noncompliance, CACE, Causal inference, Sensitivity analysis

Potential-outcomes methods such as principal stratification have gained widespread use in settings susceptible to post-treatment selection bias. In a context where treatment received can differ from treatment assigned in a manner that may depend on a key covariate, we present a method for estimating the complier average causal effect (CACE) using a Gibbs sampler for computation. We apply this method in a motivating example comparing treatments for mandible fractures in which surgeons' ability to "overrule" random assignment may have landed more severely injured patients in the more invasive treatment arm. To obtain the CACE, we model the mechanism giving rise to treatment received as possibly depending on injury severity. Varying this dependence across a feasible range produces a useful sensitivity analysis of the role of injury severity in estimation of treatment effects.

Assessing the Surrogate and Predictive Value of a Biomarker Measured After Randomization

✱Julian Wolfson, University of Washington, *julianw@u.washington.edu*; Peter B. Gilbert, Fred Hutchinson Cancer Research Center

Key Words: surrogate evaluation, principal stratification, causal inference, sensitivity analysis, vaccine trial

Surrogate assessment is complicated in the vaccine trial context by the fact that biomarkers (usually immune responses) are measured some time T after randomization to vaccine or placebo. Previous work on this problem has generally assumed no effect of the vaccine on the clinical outcome prior to T. We propose a principal stratification-based approach to quantify the surrogate and predictive value of a biomarker that relaxes this assumption. We show that the logical estimand for assessing surrogate value in this framework is a function of three non-identifiable terms. We also define a closely-related estimand which summarizes the predictive value of a biomarker and is identifiable from the observed data. A sensitivity analysis approach to estimation is taken employing the augmented trial design of Follmann, wherein both baseline predictors and close-out placebo vaccination are used.

Potential Outcomes Applied to an Intentional Weight Loss Application

✱Thidaporn Supapakorn, Kansas State University, Department of Statistics, Dickens Hall, Manhattan, KS 66506, *mmeeii@ksu.edu*; Gary L. Gadbury, Kansas State University; Christopher S. Coffey, The University of Alabama at Birmingham; Scott W. Keith, The University of Alabama at Birmingham; David B. Allison, The University of Alabama at Birmingham

Key Words: causality, obesity, latent variable, matched-pairs

Studies that explore the link between weight loss among obese individuals and mortality have met with mixed results. One possible explanation is that total weight loss may have contributions from intentional weight loss (IWL) and unintentional weight loss (UWL). These contributions may be in opposite directions from one another, yet UWL and IWL are not separately observable for subjects intending to lose weight. We focus on IWL as an unobservable latent variable. Potential outcomes are used to bound parameters associated with the distribution of IWL under an assumption of random assignment to intention. These bounds help to separate effects of IWL from UWL in a mice data example using a two sample completely randomized design. A particular matched-pairs design is also illustrated that yields additional information about variance of IWL in a population.

146 Multivariate Methods for Analysis of Genetic Data

Biometrics Section, Section on Nonparametric Statistics, Biopharmaceutical Section, IMS

Monday, August 4, 10:30 a.m.–12:20 p.m.

Inference on Low-Rank Data Matrices with Applications to Microarray Data

✱Xingdong Feng, University of Illinois at Urbana-Champaign, 725, South Wright Street, Department of Statistics, Champaign, IL 61820, *xfeng2@uiuc.edu*

Key Words: Microarray, Hypothesis Test

Probe-level microarray data are usually stored in matrices, where the row and column correspond to array and probe, respectively. Scientists routinely summarize each array by a single index as the expression level of each probe-set (gene). We examine the adequacy of a uni-dimensional summary for characterizing the data matrix of each probe-set. To do so, we propose a low-rank matrix model for the probe-level intensities, and develop a useful framework for testing the adequacy of uni-dimensionality against targeted alternatives. This is an interesting statistical problem where inference has to be made based on one data matrix. Applications of the proposed tests to GeneChip data show that evidence against a uni-dimensional model is often indicative of practically relevant features of a probe-set. The work is performed under the direction of Professor Xuming He.

Application of Multiple Correspondence Analysis and Cluster Analysis in Phenotypically Subtyping Medical and Psychiatric Disorders

✱Grace Chan, University of Connecticut, 263 Farmington Avenue, Farmington, CT 06030-2103, *chan@psychiatry.uchc.edu*

Key Words: Phenotype, Multiple Correspondence Analysis, Cluster Analysis, Medical and Psychiatric Disorder Subtype

In this talk, a procedure that employs both multiple correspondence and cluster analyses to classify finer subtypes of medical and psychiatric disorders, will be described. This kind of subtyping is valuable in both medical researches and clinical practices. For example, in genetic studies, it enables researchers to identify genetic heterogeneity in complex disorders prior to linkage analysis. Also, it enables physicians and psychiatrists to prescribe treatment targeting specific subtype of certain disorder. This talk will include an application of this procedure to opioid dependence disorder.

An FDR-Controlling Procedure for Analyzing Replicated Microarray Time Course Data with the Multivariate Empirical Bayes Statistic

✱Nancy N. Wang, University of California, Berkeley, Division of Biostatistics, 367 Evans Hall #3860, Berkeley, CA 94720-3860, *ncwang@stat.berkeley.edu*; Terence P. Speed, University of California, Berkeley

Key Words: gene expression, longitudinal data, Hotelling T^2, moderation, shrinkage

A multivariate empirical Bayes statistic proposed by Tai and Speed has been shown useful for ranking genes based on differential expression in replicated microarray time course experiments. This is known as the MB-statistic, or the moderated Hotelling T^2 statistic when all genes share a common sample size. We propose an FDR-controlling procedure for this statistic. A null distribution is obtained using the parametric bootstrap. Critical values are determined according to the empirical Bayes FDR procedure. We compared this method, through simulations, to the frequentist FDR procedure, which requires a theoretical null distribution for calculating the nominal p-values. Although our method is slightly anti-conservative; it is more robust to the variability in the estimates of the hyper-parameters, when the degree of moderation is small.

Outliers When Clustering Microarray Data

✱Johanna Hardin, Pomona College, 610 N. College ave., Dept of Mathematics, Claremont, CA 91711, *jo.hardin@pomona.edu*

Key Words: clustering, outlier detection, microarray, distance metric

Microarray data are well known to be noisy and rife with outliers. The outliers are sometimes interesting in their own right, but often they are simply poor quality measurements that should be removed from the analysis. Unlike many other statistical techniques, clustering methods will always give cluster outputs regardless of the structure of the data. Though clustering results can be enormously informative, the results can also be misleading if the data have outlying values. In particular, when clustering genes with only tens of samples, a few outlying values can easily change the direction of the relationship between a pair of genes. We propose a robust clustering technique that captures relationships between genes that are co-regulated while down-weighting relationships that are primarily due to outliers.

Reproducibility of Classification Rules Based on a Bootstrap Resampling Approach

✱Chin-Yuan Liang, The Ohio State University, Department of Statistics, Cockins Hall, Room 404, 1958 Neil Avenue, Columbus, OH 43210, *liang.88@osu.edu*

Key Words: reproducible, K-TSP, classification, Microarray, algorithm

With the advent of Microarray data, numerous methods for classification of biological samples have been proposed. However, reproducibility, or consistency of rules used in classifying observations in different samples is seldom achieved. In this paper, we propose a bootstrap resampling method to achieve a set of reproducible rules. More specifically, as an example, we first utilize the K-Top Score Pairs algorithm (K-TSP) to construct classification rules. Then, with the help of a proper rank aggregation method, we obtain rules that are deemed to be more producible. Further, because of comparability of ranks, our method can be applicable to data from different platforms. We illustrate our approach using three lung cancer data from two different platforms. Finally, it is worth noting that although our illustration is based on K-TSP, other classification methods can be utilized as well.

Sliced Inverse Factor Analysis

✱Ronghua Luo, Peking University, Changchunxinyuan 2-365, Peking University, Beijing, P.R.China, International 100871 P.R. China, *luoronghua@gsm.pku.edu.cn*; Hansheng Wang, Peking University; Chih-Ling Tsai, University of California, Davis

Key Words: BIC, Factor analysis, Sliced inverse regression

We employ the sliced inverse approach in conjunction with the factor analysis model to obtain a sufficient dimension reduction method, named sliced inverse factor analysis (SIFA). This method is particularly useful when the number of predictors is substantially larger than the sample size. We show theoretically that SIFA is able to recover the latent factor accurately. In addition, a consistent BIC-type selection criterion is developed to select the structure dimension. Both simulations and a microarray example are presented to illustrate the usefulness of the new method.

A Novel Multistage Approach To Classify Samples Across Multiple Gene Expression Experiments

✱Heather A. Adams, University of Illinois at Urbana-Champaign, 1207 West Gregory Drive, Urbana, IL 61801, *hadams1@uiuc.edu*; Sandra L. Rodriguez-Zas, University of Illinois at Urbana-Champaign; Bruce Southey, University of Illinois at Urbana-Champaign

Key Words: classification, discrimination, mixed model, gene expression experiments

A two-stage method to identify genes that classify samples across microarray experiments is presented. In the first stage, a non-linear model is used to identify associations between genes and classes within experiment. Three variations of the second stage were used to classify samples into two behaviors across experiments. The MARG approach analyzed association estimates (regression coefficient) between the expression of individual genes and odds ratio of behavior. The COND approach analyzed association estimates between individual genes and odds ratio of behavior, conditional on other genes. The JOIN approach analyzed the difference between predicted probabilities of behavior based on multiple genes. The JOIN approach identified more classifier genes, followed by COND and MARG. The two-stage approach is a valuable alternative to identify genes that consistently classify samples.

147 Statistical Inference in Biopharmaceutical Research ●

Biopharmaceutical Section, Biometrics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

Cochran-Mantel-Haenszel Weighted Miettinen and Nurminen Method for Confidence Intervals of the Difference in Binomial Proportions from Stratified 2x2 Samples

✱Kaifeng Lu, Merck & Co., Inc., Piscataway, NJ 08854, *kaifeng_lu@merck.com*

Key Words: Cochran-Mantel-Haenszel weights, Score method for variance

The Miettinen & Nurminen (1985) method is often used for constructing confidence intervals of the difference in binomial proportions from stratified 2x2 samples. However, the weighting strategy proposed in the their paper requires an iterative procedure to implement, which is nested within another iterative procedure for finding the confidence limits. This paper examines the Cochran-Mantel-Haenszel (CMH) weighting strategy that does not require an iterative procedure to implement. Compared to the weighting strategy proposed in their original paper, the new weighting strategy does not affect the performance of the resulting confidence intervals while significantly reducing the computational time.

A Nonparametric Method for Group Comparison Using Data from Two Sequential Assays with Different Lower Limits of Quantification (LLOQ)

✱Lixin Han, Wyeth, 02140, *hanl2@wyeth.com*; Jared Christensen, Wyeth; Fred Immermann, Wyeth; Ron Menton, Wyeth

Key Words: assay, preclinical, nonparametric, messy data

A problem arises from a bioassay that has a very narrow range of quantification. For assay values of different subjects to be comparable, the amount of tissue used for the assay must be the same for all subjects. This results in some subjects having assay values below the LLOQ of the assay. However, if the amount of tissue is increased to rerun the assay for subjects who have values lower than LLOQ in the first assay, the assay values are usually within the limits of quantification of the second assay. Because the assay values from the second assay cannot be easily converted to the scale of the first assay, a nonparametric method is developed to use data of both assays to compare treatment groups. The advantage of this method is demonstrated in simulation by comparisons to the left-censoring method and the method of substituting the LLOQ values by a fixed value such as LLOQ/2.

Direct Approaches to Shelf Life Estimation

✱Michelle Quinlan, University of Nebraska-Lincoln, *mquinlan22@yahoo.com*; Walter Stroup, University of Nebraska-Lincoln; James Schwenke, Boehringer Ingelheim Pharmaceuticals, Inc.

Key Words: calibration, quantile regression, shelf life, out-of-specification, random effect

The goal of shelf life estimation is to determine the time when the probability of an out-of-specification product exceeds a. Current practices approach this indirectly via interval estimation of the mean. It is unclear if this approach captures the target probability. Alternatively, one might consider direct methods. Two methods considered are calibration and quantile regression. Calibration uses a model Y=f(time) and focuses on the interval on the time axis where the predicted Y intersects the specification limit. Quantile regression changes the model's focus from the expected value to the ath percentile. Thus, the estimated shelf life is the time when the estimated quantile intersects the specification limit. Calibration and quantile regression are presented for estimating shelf life from multi-batch stability data where batch effects may be random.

The Tau-Path Test for Detecting Subpopulation Association

✱Li Yu, The Ohio State University, 404 Cockins Hall, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210-1247, *liyu@stat.osu.edu*; Joseph Verducci, The Ohio State University

Key Words: subpopulation monotone association, Kendall's tau decomposition, tau-path test, partial association

The "tau-path" test method, which is sensitive to association in an unknown subpopulation, is introduced. It is based on a decomposition of Kendall's coefficient of monotone association that sequentially identifies the strongest associations among increasing subsets of observations. The test statistic is derived from a semiparametric model for the sampling distribution of the empirical copula. A computationally efficient procedure, the component sums algorithm, is used to calculate the test statistic. Simulation studies with multivariate normal mixtures show that the proposed test is more powerful than the Mann-Kendall test against alternatives where high association is limited to a subpopulation; moreover it automatically finds the likely associated subsets. We apply the test to the NCI-60 database to discover associations between micro-RNA expression and sensitivity to drug compounds.

Differential Expression Analysis with Missing Data: Application in Two-Dimensional Gel Electrophoresis (2DGE) Analysis

✱Dan Li, University of California, Davis, 6600 Orchard Park Cir. Apt 6024, Graduate Group of Applied Mathematics, Davis, CA 95616, *dali@ucdavis.edu*; Susanne R. Berglund, University of California Davis Cancer Center; Alison Santana, University of California Davis Cancer Center; Zelanna Goldberg, University of California Davis Cancer Center; David M. Rocke, University of California, Davis

Key Words: Missing Data, Differential Expression, EM Algorithm, Two-Dimensional Gel Electrophoresis

The objective of proteomics analysis is to identify differentially expressed proteins across different experimental conditions. Often this must be done where a significant proportion of data is missing. An example of such a challenge is the missing data problem of 2DGE. The missing values can be either generated by random experimental variations or as a result of being below the limit of detection and it is not always possible to say which one is the cause of missing data. We develop a novel statistical technique that uses the Expectation-Maximization (EM) based parameter estimates and the F-test to identify the differential expression in the presence of missing data. Simulation studies reveal that our approach has the better performance compared with two other ordinary treatments of missing data problem, discarding missing data and filling in the missing values by the detection limit.

Statistical Models and Testing Methods in Clinical Trials for the Prevention of Type 2-Diabetes Mellitus

✱Xiaodan Wei, sanofi-aventis, BX2-416B, 200 Crossing Blvd, Bridgewater, NJ 08807, *xiaodan.wei@sanofi-aventis.com*; Peng-liang Zhao, sanofi-aventis; Ji Zhang, sanofi-aventis

Key Words: Diabetes prevention, Clinical trial, Impaired glucose tolerance, Hypothesis

Currently, many clinical trials are conducted in patients with impaired glucose tolerance or impaired fasting glucose to establish the prevention effect of the drug on the onset of type 2 diabetes. To clearly evaluate the prevention effect of the study medication, such trials have a washout period after the patients stop the study medication. The estimation of prevention effect would be biased if one simply compares the event rates of the two groups during the washout period due to the different dropout rates. In this paper, we first set up a hypothesis to test the prevention effect, and then provide test statistics. We also build the patient recruiting and drop out in the model so that it can be applied to both fixed study duration design and fixed stopping time design. The testing performance will be evaluated by simulation results.

A Latent Normal Approach for Equivalence Trials with Ordinal Responses Using Pairwise Likelihood

✱Juanmei Liu, University of California, Los Angeles/Kaiser Permanente, *jliuuanm@ucla.edu*

Key Words: pairwise likelihood, latent trait

Equivalence/non-inferiority trials, which attempt to demonstrate that the responses from two treatments differ by an amount that is clinically

insignificant, has become an essential part in drug/treatment assessment. In many situations, the responses of interest are usually recorded in ordinal scales. This paper demonstrates a latent variable approach with pairwise likelihood in the presence of ordered categorical data. The proposed methodology was based on the assumption that the observed ordinal variable has an underlying normally distributed trait. A real medical example is presented and compared with current results.

148 Early Human Trials and Crossover Designs ●▲

Biopharmaceutical Section, Biometrics Section

Monday, August 4, 10:30 a.m.–12:20 p.m.

Statistical Issues for First-in-Human Designs

✱Xiaoni Liu, Bristol-Myers Squibb Company, Route 206 & Province Line Road, Princeton, NJ 08543, *xiaoni.liu@bms.com*; Robert Smith, Bristol-Myers Squibb Company

Key Words: First-in-Human (FiH) study, Alternating Panel Crossover Design, Parallel Design, Sequential Panel Crossover Design, Adverse Event Rate

The main objective of Phase I First-in-Human study is to evaluate the safety and tolerability of single doses of a new drug. As a simple and safe design, the parallel design is applied in FiH studies more widely than other possible designs. On the other hand, crossover designs are superior to parallel designs in assessing the PK and PD properties of the new drug. The bias in estimating AE rates due to dropouts is the biggest concern of the crossover designs. Some simulations were applied to quantitatively measure and compare the bias of the estimated AE rates from the dropouts among the five possible FiH designs. The simulation results suggest any bias in estimating AE rates for mild and moderate dropout rates seems to be small and not to differ among these five possible FiH designs.

Linear Models for the Analysis of Alternating Panel Rising Dose Designs

Bo Jin, Merck & Co., Inc.; ✱Peng Sun, Merck & Co., Inc., *peng_sun@merck.com*

Key Words: alternating panel, rising dose, first-in-man, design

An alternating panel rising dose design (APRDD) is often used in the pharmaceutical industry in a first-in-man (FIM) clinical trial with a new drug. This paper addresses the problem of the choice of statistical models for analyzing data from alternating panel rising dose designs. Two new models are presented for pharmacodynamics (PD) data and pharmacokinetics (PK) data respectively. Comparing to a previously proposed model, the new models are shown to better reflect the study design and estimable functions.

Statistical Analysis Methods and Clinical Pharmacology Understanding on Bioavailability Comparison in Complete and Incomplete Crossover Design and Parallel Design Studies

✱Yonghua Wang, Bristol-Myers Squibb Company, *yonghua.wang@bms.com*

Key Words: crossover, study design, statistical methods, bioavailability

Crossover study design is commonly used in bioavailability and bioequivalent studies. The statistical methods were advanced since the FDA's guideline. But the new methods for the mixed effect models are constantly misunderstood and/or misused in the pharmaceutical industry. There is also misunderstanding among clinical pharmacology scientists about the bioavailability comparison in crossover study designs. The purpose of this presentation is to explain the differences among the statistical methods used in bioavailability comparisons for different kind of study designs including complete crossover designs and incomplete crossover designs as well as parallel designs and historic comparison. Challenges on explanation of adjusted geometric means which reverses the relationship of the pure geometric means from the analysis of incomplete crossover data will be explored.

The Influence of Time on Individual Response Variability in a Two-Treatment, Three-Periods Crossover Design

✱Edwin Ndum, Kansas State University, 1701 Hillcrest Dr, Apt W10, Manhattan, KS 66502, *ndum@ksu.edu*; Gary L. Gadbury, Kansas State University

In medical studies, treatment main effect is often expressed as "average difference." Based on this, a superior treatment effect may not be overall when there is substantial "subject-treatment interaction." A parameter quantifying subject-treatment interaction is inestimable in two sample completely randomized designs. Cross-over designs have been proposed as solution. However, observed individual variability may include individual treatment variability plus inherent variability over time. We propose a variable that defines a true individual effect of treatment in a three-period design and show it is not directly observable. The design compares two treatments with an estimable mean treatment effect but with a variance of individual treatment effects that is "time-biased." Certain design considerations will bound these time effects, separating the time from treatment variability.

Use of Baseline Measurements in the 2 X 2 Crossover Trial

✱Lingling Han, Merck & Co., Inc., 351 N. Sumneytown Pike, North Wales, PA 19454, *lingling_han@merck.com*; John Stufken, The University of Georgia

Key Words: crossover, baseline, repeated measurements

Crossover designs have been applied extensively in pharmaceutical industrial. It is not uncommon that measurements are available at the beginning of each period, how to incorporate the baseline measurements appropriately in the analysis under different scenarios is the question we intend to answer in this talk. Four different methods are considered in the 2 x 2 crossover design for both single measurements and repeated measurements. Analytical expressions of variances of the estimators of the treatment contrast from those methods are derived and compared under different scenarios. Simulation studies are conducted to evaluate the performance of each method. For the case of repeated measurements, different types of baselines and different assumptions for the random error terms are also considered. The methods are applied to real data analysis.

149 Section on Survey Research Methods Contributed Poster Presentations

Section on Survey Research Methods
Monday, August 4, 10:30 a.m.–12:20 p.m.

ASA Committee Web Sites

✱Alan R. Tupek, Arbitron Inc., 9705 Patuxent Woods Dr, Columbia, MD 21046, *alan.tupek@arbitron.com*; Chet Bowie, National Opinion Research Center

Key Words: Committees, Website, stakeholders, survey

In 2007, the ASA Committee on Committees formed a subcommittee to review the functionality of the ASA committee web sites. This poster session describes the process used by subcommittee. Stakeholders were identified. An assessment was made of current usage of committee web sites. A survey of a sample of committee chairs was conducted. The findings from the studies along with recommendations to improve the web site will be included in the poster session.

Application of Confirmatory Factor Analysis To Establish the Validity of a Practice-Analysis Survey

✱Heibatollah Baghi, George Mason University, 4400 University Drive, MS 5B7, Fairfax, VA 22030, *hbaghi@gmu.edu*; Teresa Panniers, George Mason University; Mary Smolenski, American Nurses Credentialing Center

Key Words: Pracitce analysis survey, Empirical job analysis, Factor analysis, Ambulatory care certification, LISREL

This study explicates practice analysis methods using ambulatory care nursing as an exemplar. Data derived from a focus group technique were used to develop a survey that was completed by 499 ambulatory care nurses. The construct validity of the survey instrument was assessed using confirmatory factor analysis with LISREL. The focus group developed a questionnaire that produced 34 knowledge statements delineating ambulatory care nursing practice. The previous study produced five factors as initial patient assessment, professional nursing issues and standards, client care management skills, technical/clinical skills, and system administrative operations. To assess the adequacy of the factors, we conducted confirmatory factor analyses on a five-factor model (Joreskog and Sorbom, 1992). Fit indices suggested that the five-factor model provided a satisfactory fit for the data.

Testing the Proportional Odds Assumption for Complex Data

✱George G. Brown, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709, *ggbrown@rti.org*

Key Words: Proportional Odds Assumption, Multinomial Logit Models, correlated data, survey data, SUDAAN

Modeling an ordinal categorical response variable is often performed using a cumulative logit link function. The proportional odds assumption (McCullagh, 1980) is key to the usage of this model. A likelihood-based version of this test is available in many statistical software packages. However, these available tests are not capable of taking into account a complex design that can include stratification, clustering, without replacement sampling and weighting or conducting the test when inference is based on replicate methods, such as Replicate Weight Jackknife and Balanced Repeated Replication. SUDAAN Release 10 will be capable of performing the test for proportional odds when any of the previously mentioned aspects are present in the data. This talk will discuss the theoretical implications of performing the proportional odds test and will demonstrate this test using survey data.

Fast Record Linkage of Very Large Files in Support of Decennial and Administrative Records Projects

✱William E. Yancey, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *william.e.yancey@census.gov*; William E. Winkler, U.S. Census Bureau; Edward H. Porter, U.S. Census Bureau

Key Words: entity resolution, approximate string search, classification

In this era of increased computing power, there is an increased need to find duplicates in very large files. This paper details the current version of Bigmatch software (Yancey and Winkler 2007) that is sufficiently fast for processing 10^17 (=300 million x 300 million pairs) for the U.S. Decennial Census and even larger administrative-record situations with billions of records. The software, via a nontrivial application of a set of blocking strategies, is known to find more than 99.5% of true matches with very small error of less than 0.5% (Winkler 2004, 1995). The software is approximately 80 times as fast as commercial Vality suite software from IBM. It is 40-50 times as fast as recent parallel software (Kim and Lee 2007; Kawai, Garcia-Molina, Benjelloun, Menestrina, Whang and Gong 2006).

Bid Design and Its Influence on Stated Willingness To Pay and Participation in a Deposit Refund Program in a Survey of Nonrefillable Plastic Pesticide Users

✱Danna L. Moore, Washinton State University, Social & Economic Sciences Research Center, Washington State University, Room 133 Wilson Hall, Pullman, WA 99164-4014, *moored@wsu.edu*

Key Words: Survey, Willingness to pay, contingent valuation, recycling, environment, bias

Currently there are no federal level rules or mandates addressing or regulating the removal of plastic from the environment as a bi-product of global pesticide manufacturers releasing non-refillable plastic containers into U.S. agriculture. Degradation of the environment from non-refillable plastic pesticide containers could have a lasting affect on agricultural land and the economic sectors that rely on healthy land for the production of food and fiber for their income generation. Adding to the complexity of recycling plastic is also the need for controlling that the plastic that is recycled is clean an free of toxic pesticide residuals in containers. Contingent Valuation (CV) is used to estimate the monetary value of a proposed change to the provision of products to pesticide licensees.

Imputation Methods for Missing Data in a Longitudinal Family Study

✱Michael D. Larsen, Iowa State University, 220 Snedecor Hall, Department of Statistics, Ames, IA 50011, *larsen@iastate.edu*; Frederick O. Lorenz, Iowa State University

Key Words: Fractional imputation, Multiple imputation, Social statistics, Structural equation models, Panel study, Psychology

Missing data are a serious problem for research in the behavioral, biological, economic, health, and social sciences. When respondents do not provide the desired data, subsequent analyses can be biased and less precise. Ignoring the fact that some observations are missing and reporting results based on those that are observed, although expedient, is not a recommended procedure. In longitudinal studies in which data are collected on subjects over time, the loss of statistical power due to missing information can become severe. Methods of multiple imputation are explored for use with the Iowa

Family Transitions Study. The study is a multi-year multi-informant data collection effort concerned with financial hardship, emotional well-being, and martial and other relationships. Effective methods for missing data are critical for maintaining statistical power and effective sample size.

Imputation for Missing Physiological and Health Measurement Data: Tests and Applications

✱Matt Jans, The University of Michigan, Michigan Program in Survey Methodology, Institute for Social Research, 426 Thomps Street, RM 4050, Ann Arbor, MI 48104, *mattjans@isr.umich.edu*; Steven G. Heeringa, Institute for Social Research

Key Words: health surveys, missing data, imputation, biological measures

We address alternative approaches to imputation for univariate estimates and multivariate regression analyses involving physiological and health measurements. From the National Health and Nutrition Examination Survey (NHANES) 2003–2004 data file we selected 5041 respondents age 20+ who have questionnaire or MEC data. Following the work of Schafer and others on NHANES III, several outcome measures were chosen for imputation (blood pressure, cholesterol, weight, general health condition, and income) We examine the results of various imputation methods for imputing these values for all cases (including MEC nonrespondents) and for item nonresponse given MEC response. Missing data rates for our variables range from 7–13% and different rates are seen across sex and age subclasses. Results from single, hot deck, regression, and multiple imputation will be compared and implications discussed.

Model-Assisted Hot-Deck Imputation

Hongjian Yu, University of California, Los Angeles; ✱Winnie Huang, University of California, Los Angeles, Center for Health Policy Research, 10960 Wilshire Blvd, #1550, Los Angeles, CA 90024, *yiichieh@ucla.edu*; Jenny Chia, University of California, Los Angeles; Sunghee Lee, University of California, Los Angeles

Key Words: Imputation, Model Assist, propensity score

We compare soft-boundary hot-deck (SBHD) imputation with model-assist hot-deck (MAHD) using simulation method. In SBHD, donors and receivers are matched in cells defined by a list of variables. There is no control on the number of observations in each cell and the number of the variables used to define the cells is limited. If at least one cell runs out donors, the last variable on the list will be dropped to order to enlarge the donor pool. In MAHD, the donor-receiver grouping is defined through the intervals and percentiles of predicted values by a model. It greatly reduces or eliminates those limitations of SBHD. Therefore it provides flexibility to optimize the procedure and possibility to produce better results. We also study grouping which utilizing propensity score of missing.

Imputing and Jackknifing Scrambled Responses

✱Inderjit S. Grewal, Punjab Agricultural University, Department of Mathematics, Statistics and Physics,, Ludhiana (PUNJAB), International 141004 India, *isg1969@yahoo.com*; Jong M. Kim, University of Minnesota, Morris; Sarjinder Singh, The University of Texas at Brownsville and Texas Southmost College

Key Words: Estimation of mean, Estimation of variance, Jackknifing, Sensitive variables

“Scrambled data are as good as scrambled eggs.” In the present investigation, it has been shown that scrambled responses on sensitive variables such as income, drugs used, induced abortions, etc. can also be imputed by following Singh, Joarder and King (1996) and can be jackknifed to estimate the variance of the resultant ratio type estimator by following Rao and Sitter (1995). Results have been simulated under different levels of untruthful reporting by following Singh, Joarder, and King (1996) and are compared with those from Rao and Sitter (1995) study.

A New Approach to Estimation of Response Probabilities When Missing Data Are Not Missing at Random

✱Michael Sverchkov, BAE Systems IT/Bureau of Labor Statistics, 2 Massachusetts Aveneu NE, Suite 1950, OSMR, Washington, DC 20212, *Sverchkov.Michael@bls.gov*

Key Words: sample distribution, complement-sample distribution, prediction under informative sampling or non-response, estimating equations, missing information principle

Often the probability of responding depends directly on the outcome value. This case can be treated by postulating a parametric model for the distribution of the outcomes before nonresponse and a model for the response mechanism. The two models define a parametric model for the joint distribution of the outcomes and response indicators, and therefore the parameters of these models can be estimated by maximization of the likelihood corresponding to this distribution. Modeling the distribution of the outcomes before nonresponse, however, can be problematic since no data is available from this distribution. We propose an alternative approach that allows estimation of the parameters of the response model by first estimating the outcomes distribution of the respondents, and then solving an estimating equation defined by the census likelihood of the response indicators.

Handling Attrition Bias on Longitudinal Studies of Mental Health Within Older Adults

✱Moh Yin Chang, University of Nebraska-Lincoln, Gallup Research Center, 200 North 11th Street, Lincoln, NE 68508, *mohyin@gmail.com*

Key Words: attrition bias, mental health, aging, sensitivity analysis, tolerence model

Attrition bias is a problematic issue longitudinal study of older adults. Attrition of the sick, impaired, and depressed respondents undermines the value of the study. Researchers attempt to adjust for health-related factors in the analysis in order to account for health-related attrition bias. Nevertheless, it is important to perform sensitivity analysis to examine the randomness of attrition given the observed information. Repeated cross-sectional studies are useful for external evaluation of the reliability of the longitudinal inferences. This study compares sensitivity analyses using selection and pattern-mixture models, and proposes methods for adjusting the attrition bias. The methods are demonstrated on a study of the impact of subjective health, chronic condition, and impairment on mental health, using data from the Health and Retirement Study.

Analysis of Using the SERVQUAL Model for Student Faculty Course Evaluations

✱Dennis F.X. Mathaisel, Babson College, Babson Hall, Babson Park, MA 02457-0310, *Mathaisel@Babson.edu*

Key Words: SERVQUAL, Student Survey

The purpose of the research was to ascertain the effectiveness of using the SERVQUAL Model (Parasuraman et al 1988) to assess the differences between college student expectations going into a course at the beginning of the semester and their final perceptions at the end of the semester. Based on past research in employing the SERVQUAL Model in other service industries, such as banking, a “gap” usually occurs because expectations generally exceed perceptions (Schneider and Bowen 1985). The author will also examine the advantages and disadvantages of the SERVQUAL technique and offer suggestions on how to improve the technique’s effectiveness.

A Quality-Control Approach for Statistical Computer Programs

✱Darryl V. Creel, RTI International, 6110 Executive Blvd., Suite 902, Rockville, MD 20852, *dcreel@rti.org*

Key Words: Statistical Programming, Quality

There are documents that provide very good high-level guidance for different aspects of the survey process. This high-level guidance is typically implemented by computer programs. This poster provides a quality approach of how to translate high-level guidance into well-documented and understandable statistical computer programs. The quality approach consists of three interrelated documents: (1) a text file that contains the executive summary of the program and a detailed section for each of the major steps of the program; (2) a flow chart containing the major steps and how they are related; and (3) computer programming techniques to facilitate the understanding and quality control (e.g., variable naming conventions, commenting, and program structure that results in a self-documenting statistical computer program).

Measurement of Perceived Health Status: Respondent and Mode Effects in a National Health Survey

✱Steven R. Machlin, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *steve.machlin@ahrq.gov*; Karen Beauregard, Agency for Healthcare Research and Quality; William Yu, Agency for Healthcare Research and Quality

Key Words: health status, proxy effect, mode effect, MEPS

Health surveys often contain a questionnaire item on whether sample persons' general health is excellent, very good, good, fair, or poor. The Medical Expenditure Panel Survey (MEPS) includes this type of question at each of 5 separate interview rounds. The MEPS is conducted via CAPI and one adult responds for all members in the household. The MEPS also includes a self administered questionnaire (SAQ) which is completed by sample adults around the time of the 2nd and 4th household interviews; it includes the same health status question as in the CAPI. In this paper, we assess the extent that type of respondent (self versus other family member), questionnaire mode (in-person CAPI versus self-administered paper), and characteristics of the respondent and sample person affect responses to the perceived health status question.

Cluster Size in Multilevel Models: The Impact of Sparse Data Structures on Point and Interval Estimates in Two-Level Models

✱Bethany A. Bell-Ellison, University of South Florida, Department of Educational Measurement & Research, 4202 East Fowler Ave, EDU 162, Tampa, FL 33620, *bbell@coedu.usf.edu*; John M. Ferron, University of South Florida; Jeffrey D. Kromrey, University of South Florida

Key Words: multilevel modeling, sample size, survey research

Sample size at each level is important to consider when estimating multilevel models. Although general sample size guidelines have been suggested, the nature of social science survey research (e.g., large number of level-2 units with few individuals per unit) often makes such recommendations difficult to follow. This Monte Carlo study focuses on the consequences of level-2 sparseness on the estimation of fixed and random effects coefficients in terms of model convergence and both point estimates and interval estimates as a function of the level-1 sample size, number of level-2 units, proportion of singletons (level-2 units with one observation), colinearity, intraclass correlation, and model complexity. SAS IML was used to simulate 5,000 data sets across 5,760 conditions. Results are presented in terms of statistical bias, confidence interval coverage, and rates of model non-convergence.

State and County Small-Area Estimation Using the National Health Interview Survey (NHIS) and Behavioral Risk Factors Surveillance System (BRFSS)

✱Van Parsons, National Center for Health Statistics, National Center for Health Statistics, 3311 Toledo Rd, Rm 3219, Hyattsville, MD 20782, *vparsons@cdc.gov*; Nathaniel Schenker, National Center for Health Statistics

Key Words: random-digit-dialing, R software, combine surveys

The NHIS, a national face-to-face household survey, and the BRFSS, a system of state-based RDD surveys, provide a major source of health information needed by policy makers. Estimation for small areas (e.g., for counties and states) has become important, but the NHIS' inadequate sample sizes for small areas and the BRFSS' telephone coverage bias along with high nonresponse rate make direct estimates for small areas problematic. A Bayesian model-based approach, suggested by Raghunathan et al., 2007, provided a method to combine the data from both surveys and correct for the individual deficiencies. In this paper we consider a more traditional approach of using linear mixed models to combine data from the two surveys. We focus on models that can be applied to a wide range of problems and be easily applied by data users. We provide evaluations of the Bayesian and traditional models.

Variance of the With-Replacement Sample Variance

✱Eungchun Cho, Kentucky State University, 9676 Scotch Haven Drive, Vienna, VA 22181, *ecpcho@gmail.com*; Moon J. Cho, Bureau of Labor Statistics

Key Words: Moments of finite population, Randomization variance, Sample variance, Sample with replacement

The variance of a sample taken from a finite population with replacement is an unbiased estimator of the population variance. The variance of this variance estimator can be represented as a linear combination of first four moments of the population. This paper gives a formula of the variance of the sample variance in terms of the sample size and the second and fourth moments.

The Use of Paradata for Evaluating Interviewer Training and Performance

✱William P. Mockovak, Bureau of Labor Statistics, 2 Massachusetts Ave, N.E., Rm 1950, Washington, DC 20212, *mockovak_w@bls.gov*; Randall Powers, Bureau of Labor Statistics

Key Words: paradata, interviewer training, interviewer performance

Audit files produced by CAI software present a record of the actions and entries made by interviewers as an interview is completed. These data provide objective feedback on the interviewing process, which can lead to changes in training or supervision. In this study, a software tool was developed that searched for and provided frequency counts of text that described selected interviewer actions. Survey production files were then analyzed to answer a series of questions including how interviewers chose to enter data (keyboard vs. mouse), when and how they navigated in a complex instrument, how interviewers handled edit messages when they were triggered, and how more complex sections of the instrument (e.g., tables) were handled. The advantages and disadvantages of using paradata for these purposes are discussed.

Which Incentives Work Best for Respondents in Today's RDD Surveys?

✱Barbara L. Carlson, Mathematica Policy Research, Inc., P.O. Box 2393, Princeton, NJ 08543-2393, *bcarlson@mathematica-mpr.com*; Karen CyBulski, Mathematica Policy Research, Inc.; Tom Barton, Mathematica Policy Research, Inc.

Key Words: response rates, incentives, RDD surveys

As has been widely reported in the literature, response rates for telephone surveys—particularly RDD surveys—have been declining significantly in recent years. Mathematica Policy Research (MPR) has conducted several rounds of the Community Tracking Study Household Survey (CTS) for the Center for Studying Health System Change since 1996, to inform health care decision makers about changes in the U.S. health care system, and how such changes affect people. Obtaining high response rates for the CTS has been increasingly difficult each round, and round 5 (2007–08) posed the greatest challenges yet. MPR implemented various incentive and other strategies, including varying dollar amounts and the form and timing of payments. This paper shows how these strategies, implemented within random sample replicates, affected response rates and the effort needed to obtain completed interviews.

Year-to-Year Correlation in National Health Interview Survey Estimates

✱Chris Moriarity, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *cdm7@cdc.gov*; Van Parsons, National Center for Health Statistics

Key Words: complex sample, variance estimation, pooling survey data

The National Health Interview Survey (NHIS) is a multi-purpose health survey conducted by the National Center for Health Statistics (NCHS). Public-use microdata files and complex sample design variance estimation structures are available at the NCHS Internet web site. In 2007, NCHS expanded the variance estimation structures to cover all available years of public-use microdata files, which enabled researchers to compute appropriate variance estimates when conducting analyses of NHIS annual data pooled across years. We present research findings about the level of correlation in NHIS estimates from year-to-year, which provide insights in areas such as: 1) the amount of error in variance estimates for pooled data within a sample design period if the annual data are incorrectly assumed to be uncorrelated; 2) the trend in correlation over a sample design period.

On Using Sample Survey Techniques To Collect and Analyze Data Pertaining to Vehicular Traffic and Parking Problems on a University Campus

Ray Okafor, University of Lagos; ✱Ugochukwu Mbata, University of Lagos, Department of Mathematics, Faculty of Science, Lagos, 9600001 Nigeria, *okaforray@yahoo.com*; Jeanvive Amaraegbu, University of Lagos; Adebisi Fetuga, University of Lagos; Adebayo Jawando, University of Lagos; Olabaode Fasuyi, University of Lagos; Ajayi Olalusi, University of Lagos; Babalola Apara, University of Lagos; Jamaica Tamunotonye, University of Lagos; Titilayo Smith, University

Key Words: Traffic and Parking Problems on university Campus, Sample Surveys Conducted by Students, Results, Three in Ten Students used cars

In this article, we report the conduct, analysis and results of similar but independent sample surveys conducted by a group of undergraduate students of the University of Lagos (UNILAG, for short), Lagos, Nigeria. What motivated this study is the very serious vehicular traffic and parking problems that lately have become a feature of daily life on the main campus of UNILAG. With so many vehicle entering UNILAG on a daily basis, the pertinent question has been: where are these vehicles coming from? Which group among the diverse groups constituting the UNILAG community has the largest number of vehicles? Finding answers to these and other related questions was what spurred the referenced students to use the platform of their final-year projects in 2006/2007 to apply sample survey techniques they studied in class to tackle real-world problem that directly affected them.

Hardy-Weinberg Equilibrium for Unequal-Probability Genotype Samples

✱Thomas Lumley, University of Washington, Dept of Biostatistics, Box 357232, Seattle, WA 98195-7232, *tlumley@u.washington.edu*

Key Words: genomics, biased sampling, whole-genome association

Hardy-Weinberg equilibrium (HWE), which specifies the relationships between allele frequencies under random mating, has been central to genetics for 100 years. One important application in modern times is in quality control for high-throughput genotyping, where departures from HWE are a valuable diagnostic. Some large cohort studies are conducting large-scale genotyping on a subsample of their participants, to reduce costs. HWE will typically not hold in the subsample, especially when participants experiencing medical events are oversampled. I will discuss HWE testing under subsampling. It is straightforward to develop a large-sample, design-based test, which I will present, but large-sample tests for HWE have notoriously poor performance at practical sample sizes and allele frequencies. I will also present a modification of Fisher's exact HWE test for unequal probability sampling.

Surveying Parts To Construct the Whole: Sampling and Estimation Issues

✱Eugene M. Burns, Energy Information Administration, Forrestal Building, EI-63, 1000 Independence Ave, SW, Washington, DC 20585, *eugene.burns@eia.doe.gov*; Jay Olsen, Energy Information Administration

Key Words: PPS systematic sampling, survey methodology, auxiliary information, nonresponse

One challenge faced by the Commercial Buildings Energy Consumption Survey (CBECS) has been that of surveying strip shopping malls. These malls consist of unrelated establishments sharing a common building. Respondents knowledgeable about energy and related characteristics of the entire mall are difficult to find. This presentation describes the multilevel data collection strategy adopted by the 2007 CBECS. Interviewers compiled rosters of mall tenants, which were used to draw a PPS systematic sample of establishments. Interviews were conducted with the mall manager, as well as with the selected establishments. Estimation involves developing a strategy for combining these different pieces of information to reconstruct the single mall building. Auxiliary information, from the roster and from other malls, assisted in the reconstruction; inconsistency and nonresponse were problems.

Mobile Phone Survey Methodology in China

✱Xiaoyin Sun, The University of Hong Kong, China, *xysun@hku.hk*; John Bacon-Shone, The University of Hong Kong

Key Words: survey methodology, mobile phones, survey incentives, sampling frames, crime victimization

Mobile phones penetration is increasing very fast in China, with more mobile phones than the adult population in some big cities and rapidly expanding even in rural areas with relatively low fixed line penetration. This trend suggests mobile phones will soon play a critical role in survey based research in China. This study focuses on the methodological questions that need to be answered before using mobile phone surveys in China instead of fixed line telephone surveys, including fieldwork organization, frame coverage, response rates and the impact of incentives. It also yields cross city comparisons on crime victimization.

Response Quality Among Reluctant Respondents

✱Olena Kaminska, University of Nebraska-Lincoln, 200 N. 11th Str.,, 4th floor, Lincoln, NE 68508, *kaminol@gmail.com*

Key Words: reluctant respondents, converted refusals, response quality, measurement error, European Social Survey

While increasing response rate by converting refusals aims to minimize nonresponse error, it can possibly lead to higher measurement error if reluctant respondents are less motivated to attend carefully to respondent tasks. Using data from the second round of European Social Survey conducted in 2004/2005 in Netherlands, we investigate whether reluctant respondents tend to provide lower data quality. Willing and reluctant respondents are compared on measures of acquiescence, straight responses, don't know responses, middle and extreme responses, random selection of response alternatives (coin-flipping), as well as interviewer evaluation of respondent effort. The implications of the study for the future waves of the European Social Survey are discussed.

Cell Phone–Only Research at Arbitron: Statistical Analyses

✱Richard Griffiths, Arbitron Inc., 9705 Patuxent Woods Drive, Columbia, MD 21045, *richard.griffiths@arbitron.com*

Key Words: Cell Phone Only, Noncoverage, Survey, Sampling, Weighting

As the size of the U.S. population that lives in households with no landline telephones has burgeoned, Arbitron has undertaken a plan of study aimed at finding methods to mitigate potential bias due to frame noncoverage of this population in their radio audience measurement surveys. This plan of study encompasses a number of different facets, including an examination of the cost-effectiveness of sampling the cell-phone-only (CPO) population. This paper reports the results of statistical analyses designed to study the characteristics of the CPO population and how the information gleaned from these analyses might be used to improve the quality of Arbitron's audience measurement estimates in the face of the CPO issue.

Controlling Effect of Sample Design on Principal Components Analysis: A Simulation Study

✱Francine Barrington, Macro International, *francine.barrington@macrointernational.com*; Andrey Vinokurov, Macro International; Pedro J. Saavedra, Macro International

Key Words: cluster sampling, eigenvalues, factor analysis, stratified sample, unequal weights, intraclass correlation

Principal Components Analyses are typically conducted without taking into account the sampling design. Controlling for variables that are part of the sample design may affect the interrelationship of variables in a manner that is difficult to interpret. Controlling for clusters, strata or probabilities of selection may affect the substance or the stability of the results. In order to examine this issue, a data base of ZIP code areas was stratified and clustered by counties or combinations of adjacent counties. PCAs were conducted for several random samples, for cluster samples using large clusters or small clusters, and for several stratified designs. Had the objective been to obtain results similar to those of a random sample, adjusting for sample design was found to be counterproductive. Simulations also indicated that stability of results is in part a function of sample design.

Coevolution of Multivariate Optimal Allocations and Stratum Boundaries

✱Charles Day, Internal Revenue Service, P. O. Box 2608, Washington, DC 20013, *Charles.D.Day@irs.gov*

Key Words: Stratified Sampling, Genetic Algorithm, Evolutionary Algorithm, Convex Optimization

Coevolution in evolutionary algorithms allows solutions to two interdependent optimization problems to be determined simultaneously, similar to the evolution of symbiotic species in nature. There are many methods for determining multivariate optimal allocations in stratified sampling, including the use of an evolutionary algorithm (Day, 2006). There are also widely accepted methods for determining optimal stratum boundaries. This work presents a method for simultaneous determination of multivariate optimal allocations and stratum boundaries using the concept of coevolution in evolutionary algorithms.

Accounting for Sampling Design in Complex Surveys: A Jamaican Example

✱Novie O.M. Younger, The University of the West Indies, Jamaica, P.O. Box 25580, TMRI, Kin # 3960, Miami, FL 33102, *novie.younger@uwimona.edu.jm*; Rainford Wilks, ERU-TMRI

Key Words: sample weights, clustering, sample design, complex surveys, analysis of complex survey data

There is always the question of whether to incorporate sampling design into the analysis of complex survey data. Failure to account for aspects of sampling design can lead to creation of mis-specified models and a lack of understanding of the basis for a given type of analysis may lead to spurious inference. We carry out analysis of data from the Jamaica Healthy Lifestyle Survey 2000 with and without accounting for two aspects of sampling design - clustering and weighting of observations. The analyses examine risk factor relationships with chronic disease outcomes and reveal that, for some outcomes, the inference differs depending on the aspect of sampling design that has been omitted. We give reasons for the differences in inference and suggest recommendations for when to use sample clustering and sample weights in the analysis.

Effects of Sampling and Screening Strategies in an RDD Survey

Anthony M. Roman, University of Massachusetts-Boston; ✱Rebecca Crow, University of Massachusetts-Boston, CSR, 100 Morrissey Blvd., Boston, MA 02125, *becky.sheridan@umb.edu*; Elizabeth Eggleston, Research Triangle Institute; Charles F. Turner, Queens College and Graduate Center, CUNY; Susan M. Rogers, Research Triangle Institute; Sylvia Tan, Research Triangle Institute

Key Words: RDD survey, dual frame sampling, survey costs

The Monitoring Sexually Transmitted Infections Project is an RDD telephone survey being conducted in Baltimore, Maryland from 2006 through 2009. The project screens households for people 15-35 years of age and then randomly selects one eligible person. To gain efficiency and reduce costs, three methodological changes were adopted after several months of data collection. These changes were 1) in the wording of the screening questions aimed at identifying households with someone 15-35, 2) implementing a sampling strategy that incorporated the use of lists that identified the ages of residents within some households, and 3) altering the probability of selection of an eligible respondent within a household to favor selecting someone who initially answered the telephone. This paper will examine the effects of these strategies on survey cost and sample composition.

The Impact of Income Imputation Using Cascading Partial Income Information in California Health Interview Survey

✱Yifeng J. Chia, University of California, Los Angeles, 10960 Wilshire Blvd., Center for Health Policy Research, Los Angeles, CA 90024, *ychia@ucla.edu*; Winnie Huang, University of California, Los Angeles; Hongjian Yu, University of California, Los Angeles

Key Words: Income Imputation, Cascading sequence, telephone interview, California Health Interview Survey

The collection of income information is difficult in most survey. High nonresponse rate on income causes the limitation of data use. The income questions in telephone survey are often designed in a cascading sequence of questions in which the interview would ask a cascading sequence of income questions when respondent fails to provide the exact income amount. Using this partial information of income, we can generate the income range to improve the accuracy of income imputation. The paper discusses the characteristics of respondents who provide income information in full, partial or none. The purpose of this paper is to investigate the impact on income imputation by comparing the data from the California Health Interview Survey (CHIS) and Current Population Survey (CPS).

150 Section on Statistics in Sports Speaker with Lunch (fee event)

Section on Statistics in Sports

Monday, August 4, 12:30 p.m.–1:50 p.m.

It's a Nice Hobby, but ... Making the Transition from Providing Insight into Statistics Through Sports to Providing Insight into Sports Through Statistics

✱James J. Cochran, Louisiana Tech University, College of Business, PO Box 10318, Ruston, LA 71272, *jcochran@cab.latech.edu*

Key Words: sports, education, modelling, operations research, analysis, data

Throughout my career I have been repeatedly reminded (usually politely) by Deans and Department Chairs that my work in sports is not highly valued. However, recent changes in administration at my college have led to a very happy surprise - my new Dean and Department Chair see great value in sports applications! The Louisiana Tech College of Business is developing an undergraduate minor in Sports Marketing, and I have been asked to design and teach a course in Sports Analytics as part of this new minor. I will discuss some of the ways I have used sports to teach statistics and operations research, some of the ways I intend to now use statistics and operations research to teach sports, and some of the differences I anticipate will arise as I reverse my orientation.

151 Section on Bayesian Statistical Science Roundtable with Lunch (fee event)

Section on Bayesian Statistical Science

Monday, August 4, 12:30 p.m.–1:50 p.m.

Informative Priors and Sensitivity Analysis for Missing Data and Causal Inference

✱Joseph Hogan, Brown University, Center for Statistical Sciences, 121 South Main Street, Box G-S121-7, Providence, RI 02912, *jwh@brown.edu*

Key Words: dropout, counterfactuals, treatment effects, selection bias, unmeasured confounding, incomplete data

Inferences from incomplete data are not possible without unverifiable assumptions. This is true for complicated missing data and causal inference models, standard approaches such as GEE, propensity score methods, and instrumental variables (to name just a few). Assumptions about the conditional distribution of missing data given observed data are often made for convenience, but cannot be empirically checked. In that sense, they can be viewed as strong priors on the missing data distribution. Several questions will be posed for discussion: Should MAR and ignorability be the 'industry standard'? Should external sources of information, either qualitative or quantitative, be incorporated into a model? What are the characteristics of an effective sensitivity analysis, and how should they be reported. Where should we focus new research in the area of incomplete data?

152 Biopharmaceutical Section Roundtables with Lunch (fee event)

Biopharmaceutical Section

Monday, August 4, 12:30 p.m.–1:50 p.m.

Analytical Issues and Practical Solutions for Missing Data in Clinical Trials

✱Jayawant Mandrekar, Mayo Clinic, 2665 Boulder Ridge Drive NW, Rochester, MN 55901, *mandrekar.jay@mayo.edu*

Key Words: Missing Data, Clinical Trial

Missing data is a reality in clinical trials, and the practical complexities associated with analysis of missing data need attention. A case study comprising a recently completed Phase III randomized, double-blinded, placebo-controlled trial will be used to motivate the discussion. Scientific rationale for the trial, the a priori assumptions used at the design stage, issues encountered due to unplanned proportions of missing data that led to revisions of the study design and endpoints, and the involvement of the regulatory and oversight committees will be the background for this discussion. This session is intended to identify possible and practical ways to handle analyses of such data.

Adaptive Designs: Sample Size Re-Estimation and Interim Analyses

✱Dennis King, STATKING Consulting, Inc., 759 Wessel Drive, Unit 6, Fairfield, OH 45014, *statking@statkingconsulting.com*

Key Words: Sample Size Re-estimation, Interim Analyses

We will discuss the statistical methods used in sample size re-estimation and interim analyses in industry-sponsored clinical trials. The pitfalls in implementing these procedures will be discussed.

Design and Analysis of Medical Device Clinical Studies for Purposes of FDA Marketing Approval

✱Gary Kamer, U.S. Food and Drug Administration, 1350 Piccard Drive, Suite 100, (HFZ-550), Rockville, MD 20850, *gary.kamer@fda.hhs.gov*

Key Words: study design, statistical analysis, public health, regulatory setting

Certain study designs and statistical analyses may be either unique to or more commonly used in a public health regulatory setting (such as FDA) than in other research situations. This session will provide an opportunity to identify and discuss not only these methods, but the reasons for their use. Several topics may be suggested by the roundtable leader, but discussion resulting from actual participant experiences will prove most beneficial.

How Not To Be Lost in Translation: Expanding Roles for Consultant Statisticians in the Era of Multinational Drug Development

✱Yoko Adachi, U.S. Food and Drug Administration, 7500 Standish Pl, HFV-105, Rockville, MD 20855, *yoko.adachi@fda.hhs.gov*

Key Words: Multi-Region, Consulting, International Clinical Trials, Regulatory Difference, Cultural Barriers, Biostatistics

As more clinical trials are conducted in regions other than the United States and Western Europe, various issues emerge for consultant statisticians who work with sponsors with various backgrounds in terms of nationality, resources, and drug development experiences. This discussion will focus on opportunities and challenges consultant biostatisticians face when a drug is developed and marketed, such as regulatory differences, cultural barriers, and communication issues.

Innovative Statistical Design Using Biomarkers in Early-Phase Cancer Clinical Trials

✱Ying Lu, University of California, San Francisco, Department of Radiology Box 0946, San Francisco, CA 94143-0946, *ying.lu@radiology.ucsf.edu*

Key Words: Phase I and II trials, Cancer, Biomarkers, Design, Cytostatic treatment, Multistage

Testing cytostatic cancer treatments using proper biomarkers imposes challenges to the conventional statistical designs of early phase trials. Biomarkers have been used to identify patients most likely to receive the treatment benefits and to evaluate the early sign of efficacy. New statistical designs should integrate these markers into statistical decision process, rather than treat them as the secondary or exploratory endpoints. We will discuss recent research progress and the recommendations from the task force on methodology for the development of innovative cancer therapies. A list of references will be provided.

153 Section on Government Statistics Roundtable with Lunch (fee event)

Section on Government Statistics
Monday, August 4, 12:30 p.m.–1:50 p.m.

How Should Prison Inmates Be Incorporated in Census Tabulations?

✱Tom Belin, University of California, Los Angeles, 51-267 Center for Health Sciences, Dept. of Biostatistics, Los Angeles, CA 90095-1772, *tbelin@ucla.edu*

Key Words: Redistricting, Prisoners, Group Quarters, Census Bureau, Usual Residence, New York

The U.S. Census Bureau has traditionally allocated people to places based on their "usual residence," but when a prison location is far from the former residences of the inmates, allocation based on "usual residence" can have different implications for political redistricting than would other allocation approaches. This roundtable discussion, led by a recent member of the ASA Census Advisory Committee, will discuss the controversy that has emerged around this issue and consider the merits of a "federalist" approach that would distinguish federal and state prisoners from other group-quarters residents in tabulations for redistricting purposes, which would allow jurisdictions to tailor policy decisions (including redistricting maps) to judgments about whether to include inmates.

154 Section on Health Policy Statistics Roundtable with Lunch (fee event)

Section on Health Policy Statistics
Monday, August 4, 12:30 p.m.–1:50 p.m.

Is There a Future for Surrogate Marker Evaluation in Randomized Clinical Studies?

✱Geert Molenberghs, Hasselt University, Center for Statistics, Diepenbeek, B3590 Belgium, *geert.molenberghs@uhasselt.be*

Key Words: surrogacy, meta-analysis, information theory, prediction

Surrogate endpoints are often considered instead of true endpoints in clinical studies, especially when such endpoints can be measured earlier and/or with less burden. It is not surprising that the use of surrogate endpoints in clinical practice is increasing, in spite of early skepticism. We build on the work of Prentice (1989), Freedman et al. (1992), and Buyse et al. (2000) to frame evaluation within a meta-analytic setting to overcome difficulties that arise when done in a single trial. The meta-analytic framework has been extended in many ways, and proposals have been made to unify the disparate collection of validation measures. Implications for design, analysis, and prediction have been addressed. Is there a future for surrogate marker evaluation and, if so, which steps for improvement are needed?

155 Section on Nonparametric Statistics Roundtable with Lunch (fee event)

Section on Nonparametric Statistics
Monday, August 4, 12:30 p.m.–1:50 p.m.

Nonparametric Statistics: A Look into the Future

✱Peter G. Hall, The University of Melbourne, Department of Mathematics and Statistics, Melbourne, International 3010 Australia, *p.hall@ms.unimelb.edu.au*

Key Words: nonparametric statistics, future directions

An exciting aspect of research in statistics today is the pace at which the frontiers are moving forward. In no area of our discipline are things changing faster than in nonparametric statistics. What topics lie at the forefront of modern nonparametric statistics, and what will be the key directions tomorrow? Is the statistics community continuing to set the pace, or are other fields (e.g., computer science) leading the way in important areas? Come along and find out.

156 Section on Physical and Engineering Sciences Roundtable with Lunch (fee event)

Section on Physical and Engineering Sciences
Monday, August 4, 12:30 p.m.–1:50 p.m.

Statistical Issues in Cybersecurity

✱Joanne R. Wendelberger, Los Alamos National Laboratory, Statistical Sciences Group, MS F600, Los Alamos, NM 87545, *joanne@lanl.gov*

Key Words: computing, vulnerabilities, anomaly detection, security

With the development of approaches for attacking electronic systems and information comes a need for investigating, analyzing, and preventing increasingly sophisticated cybersecurity threats. This roundtable luncheon will provide a forum for discussing statistical issues that arise in cybersecurity applications and alternative approaches for identifying and investigating potential threats through the application of statistical techniques in conjunction with computer security experts.

157 Section on Quality and Productivity Roundtables with Lunch (fee event)

Section on Quality and Productivity
Monday, August 4, 12:30 p.m.–1:50 p.m.

Nonclinical Applications in the Pharmaceutical Industry

✱Richard K. Burdick, Amgen, Inc, 4000 Nelson Road, MS AC-22A, Longmont, CO 80503, *rburdick@amgen.com*

Nonclinical applications typically concern issues related to validation, quality, and stability. Discussion at this table will focus on statistical methods used for setting process validation acceptance limits and stability studies.

Being a Female Industrial Statistician

✱Diane K. Michelson, International Sematech Manufacturing Initiative, 2706 Montopolis Dr., Austin, TX 78741, *di.michelson@ismi.sematech.org*

Key Words: women in statistics, industrial statistics, communication

At this roundtable, we will discuss both the issues and benefits of working as a female industrial statistician. The intent is not to field complaints, but to determine behaviors that can enhance the productivity of a female industrial statistician. Is there an effect of gender on the working relationship between statistician and client? If so, what can we do to either take advantage of a positive effect or mitigate a negative effect? What is different for females in industry versus females in academia or government? What about relationships with management, including promotions and performance reviews? Come with examples of workplace relationships for advice and answers.

158 Section on Statistical Education Roundtables with Lunch (fee event)

Section on Statistical Education
Monday, August 4, 12:30 p.m.–1:50 p.m.

Improving Utility and Fairness of Course Evaluations

✱Monnie McGee, Southern Methodist University, SMU P.O. Box 750332, 3225 Daniel Ave Room 144 Heroy, Dallas, TX 75275-0332, *mmcgee@smu.edu*

Key Words: Course Evaluation, Teaching, Data Analysis

Every professor is subjected to student evaluations, and every professor knows students are not really qualified to judge teaching. Students know what they like and don't like, and what affects their grades, but they do not know the difficulties of teaching a course. Furthermore, there have been many articles that suggest variables such as class size, time of day, and even the professor's gender—none of which are under the professor's control—affect the results of evaluations. Can we do something intelligent to make sure these evaluations are fair and used fairly?

How Do We Promote Quantitative Literacy Across the Undergraduate Curriculum, and How Will We Know if We Are Successful?

✱A. John Bailer, Miami University, Dept. of Math. & Stat., Oxford, OH 45056, *baileraj@muohio.edu*

Key Words: numeracy

Quantitative literacy is promoted, at least implicitly, in every statistics class we teach. We share a goal of producing students with "an aggregate of skills, knowledge, beliefs, dispositions, habits of mind, communication capabilities, and problemsolving skills that people need to engage effectively in quantitative situations arising in life and work." How do we promote quantitative literacy beyond the first statistics course so it is infused throughout the curriculum? What should each student know upon graduation if they are deemed "quantitatively literate"?

Barriers to Teaching Introductory Statistical Inference

✱Mark Berenson, Montclair State University, Dept. of Management & Information Systems, Montclair, NJ 07043, *berensonm@mail.montclair.edu*

Key Words: statistical inference, teaching barriers

Teaching introductory statistical inference is often frustrating. Various pedagogical dilemmas concerning course content need to be thoroughly considered if we want our students to complete the topic and say it was relevant and practical and that it increased their critical thinking and analytic skills. Attendees of this roundtable luncheon will address such barriers to learning inference as the obtuse statistical terminology employed, the noncohesive textbook table presentations provided, the inappropriate software algorithms offered, and the inadequate sequencing of topics selected. Responsibility for improvements must be borne by text and software publishers, text and software authors, and faculty members providing classroom instruction.

159 Section on Statistical Graphics Roundtable with Lunch (fee event)

Section on Statistical Graphics

Monday, August 4, 12:30 p.m.–1:50 p.m.

Data Visualization for the Masses

✱Diane Lambert, Google, Inc., New York,, NY 10011, *dlambert@google.com*

Key Words: exploratory data analysis

New data visualization applications such as Google Trends, Many Eyes, and Swivel make it easy for people to create and share statistical graphics. Much (most?) of this work is progressing without statisticians. Is there anything statistics can contribute to or learn from this movement?

160 Section on Statisticians in Defense and National Security Roundtable with Lunch (fee event)

Section on Statisticians in Defense and National Security

Monday, August 4, 12:30 p.m.–1:50 p.m.

Statistical Support for Public Health Preparedness and Medical Emergency Response

✱Myron Katzoff, Centers for Disease Control and Prevention, CCHIS/NCHS/ORM, 3311 Toledo Road, Hyattsville, MD 20782, *mjk5@cdc.gov*

Key Words: biosurveillance, health monitoring, outbreak detection

The issuance of Homeland Security Presidential Directive 21 in October 2007 established the need for frequent and informed discussion of issues centering on cross-agency collaborative efforts directed at a national strategy for public health and medical preparedness. This roundtable is intended to promote acquisition of the means and development of the statistical data and systems that will enhance our ability to address our homeland security problems in this area. The discussion is expected to focus on challenges to the statistical community for collaborating and information sharing and analytic methodologies and data.

161 Section on Statistics in Epidemiology Roundtable with Lunch (fee event)

Section on Statistics in Epidemiology

Monday, August 4, 12:30 p.m.–1:50 p.m.

Analysis of Longitudinal Complex Survey Data Sets

✱Punam Pahwa, University of Saskatchewan, 103 Hospital Drive, Dept. of Community Health and Epidemiology, Saskatoon, SK S7N OW8 Canada, *pup165@mail.usask.ca*

Key Words: epidemiologist, complex survey, longitudinal, GEE, robust

Many epidemiologists, social scientists, clinicians, and applied statisticians use longitudinal complex survey data for their research. It is important to conduct statistical analysis by using appropriate statistical methods. In this roundtable discussion, two approaches used to analyze the longitudinal data available from a multistage complex survey will be discussed. The first (Generalized Estimating Equations) is based on the assumption that study design involves only subject-level clustering due to repeated measurements, thus ignoring the complexities of survey design. The second incorporates complexities of the design and the subject-level clustering. Complexities of the design are incorporated via appropriate variance estimation techniques. These analytical approaches will be discussed with the help of examples.

162 Section on Survey Research Methods Roundtable with Lunch (fee event)

Section on Survey Research Methods

Monday, August 4, 12:30 p.m.–1:50 p.m.

Election Statistics Results

✱Wendy Rotz, Ernst & Young LLP, 1101 NY Avenue, Washington, DC 20005, *wendy.rotz@ey.com*

Key Words: election, audit, vote, ballot, poll, recount

Voter confidence and transparent elections are essential to modern democracies. States are legislating precinct-level, sample-based audits. New Jersey just passed legislation requiring elections be audited to achieve particular power levels to detect election altering miscounts, and Florida is working on legislation modeled after New Jersey. A statistical foundation in the development of election auditing procedures aids transparency and improves voter confidence in election results and the democratic process in both the United States and abroad. Come share ideas concerning statistical methodology in the election process. Statistical issues may include random recounts in all elections, 100% recounts in close elections, paper trails, Six Sigma applications, process control, exit polling, and more.

163 Social Statistics Section Roundtable with Lunch (fee event)

Social Statistics Section
Monday, August 4, 12:30 p.m.–1:50 p.m.

Delving into Multiple Methods To Examine the Relationship Between Psychological Measurement and Empirical Outcome

✱Gideon D. Bahn, Loyola University Chicago, 1623 Coolidge Ave., Wheaton, IL 60187, *gideonbahn@yahoo.com*; Martha E. Wynne, Loyola University Chicago; Mary Satchwell, Loyola University Chicago

Key Words: Korean American, Confucian five moral codes, academic achievement, Structural Equation Modeling, Stepwise Analysis, Post-hoc comparison

Statisticians often have to use multiple methods to answer complicated research questions. This study delves into multiple different methods to examine the relationship between Korean American students' academic achievement and their family background, focusing on the fundamental ideology of Korean education, the five Confucian moral codes. The data are analyzed through the sensitive analysis for Hybrid Path Model and Path Analysis in Structural Equation Modeling, post-hoc analysis in Analysis of Variance and Backward Method in Multiple Linear Regression Modeling. The results suggest that Korean American students' high academic achievement is related more to their parents' financial support for their children's education than to the degree to which parents retain traditional beliefs about Korean educational values embodied in the five Confucian moral codes.

164 Late-Breaking Session I: The Accuracy of Election Polls

ASA, ENAR, IMS, SSC, WNAR, Social Statistics Section, Scientific and Public Affairs Advisory Committee, Section on Survey Research Methods
Monday, August 4, 2:00 p.m.–3:50 p.m.

Evaluating the Performance of the 2008 Pre-Election Polls in the Primaries

✱Michael W. Traugott, The University of Michigan, 4230 ISR, 426 Thompson Street, Ann Arbor, MI 48106-1248, *mtrau@umich.edu*

Sources of Variation in Pre-Election Polling

✱Cliff Zukin, Rutgers, The State University of New Jersey, Public Policy Department, New Brunswick, NJ 08901, *zukin@rci.rutgers.edu*

This paper presents a variety of answers to the question of "Why Do Election Polls Differ?" It examines issues of timing, sampling, questionnaire wording, weighting and the identification of "likely voters" among others

Understanding and Communicating Sources of Measurement and Operational Error in Opinion Polls: Beyond Sampling Error

This paper will present examples of problems in accurately measuring preference in election polls, focusing on poll timing, measuring intensity, and the interaction between respondent and interviewer. Have interviewer characteristics like race, gender and age affected poll accuracy? When they report polls, are journalists giving polls more precision than they deserve?

165 Semiparametric Regression and High-Dimensional Data ●

Section on Nonparametric Statistics, IMS, Section on Physical and Engineering Sciences, WNAR
Monday, August 4, 2:00 p.m.–3:50 p.m.

Semiparametric Modeling with Applications to Powerful Testing for Gene-Environment Interactions

✱Raymond Carroll, Texas A&M University, 3143 TAMU, Department of Statistics, College Station, TX 77843-3143, *carroll@stat.tamu.edu*; Arnab Maity, Texas A&M University; Nilanjan Chatterjee, National Cancer Institute; Enno Mammen, University of Mannheim

Key Words: Repeated measures, Genetic epidemiology, Nonparametric regression, Semiparametric methods, Score testing

Motivated from the problem of testing for genetic effects on complex traits in the presence of gene-environment interaction, we develop score tests in general semiparametric regression problems that involves Tukey style 1 d.f form of interaction between parametrically and nonparametrically modeled covariates. We find that the score-test in this type of model, as recently developed by Chatterjee et al. (2007) in the fully parametric setting, is biased and requires undersmoothing to be valid in the presence of nonparametric components. Moreover, in the presence of repeated outcomes, the asymptotic distribution of the score test depends on the estimation of functions which are defined as solutions of complex integral equations, making implementation difficult and computationally taxing. We develop adjusted score statistics which are unbiased and asymptotically efficient for this problem.

Functional Embedding for High-Dimensional Data

✱Hans G. Müller, University of California, Davis, Department of Statistics, One Shields Avenue, Davis, CA 95616, *mueller@wald.ucdavis.edu*

Key Words: Functional Data Analysis, Dimension Reduction, Classification, Survival Prediction, Gene Expression, Univariate Scaling

Functional embedding aims at dimension reduction for large p, small n data by ordering the elements of the p-dimensional predictor vectors via univariate scaling, applied to correlation-based dissimilarities. Under regularity conditions, the embedding maps predictor vectors to a smooth random function for each subject, subsequently represented by just a few functional principal component scores. For high-dimensional gene expression data as predictors, functional embedding coupled with functional regression procedures outperforms previous methods for the classification of leukemia and for the prediction of survival time of lymphoma patients. Talk based on collaborations with Ping-Shi Wu, as well as with Kun Chen and Jane-Ling Wang.

Assist: A Suite of S Functions Implementing Spline-Smoothing Techniques

✱Yuedong Wang, University of California, Santa Barbara, Department of Statistics and Applied Probability, Santa Barbara, CA 93106, *yuedong@pstat.ucsb.edu*; Chunlei Ke, Amgen, Inc

Key Words: smoothing spline, R package, reproducing kernel Hilbert spaces, semi-parametric regression, non-parametric nonlinear regression, random effects

We present a suite of user friendly R functions for fitting various smoothing spline models including (a) nonparametric regression models for independent and correlated Gaussian data, and for independent binomial, Poisson and Gamma data; (b) semiparametric linear mixed-effects models; (c) nonparametric nonlinear regression models; (d) semiparametric nonlinear regression models; and (e) semiparametric nonlinear mixed-effects models. The general form of smoothing splines based on reproducing kernel Hilbert spaces is used to model nonparametric functions. Thus, these R functions deal with many situations in a unified fashion. Well-known special cases are polynomial, periodic, spherical, thin-plate and L splines, GAM, SS ANOVA, projection pursuit, multiple index, varying coefficient, functional linear and self-modeling nonlinear regression models.

Bayesian Smoothness and Shrinkage Priors in Regression

✱Ludwig Fahrmeir, University of Munich, Germany, Ludwig-Maximilians-University, Ludwigstrasse 33, Munich, International D-81375 Germany, *ludwig.fahrmeir@stat.uni-muenchen.de*; Thomas Kneib, University of Munich, Germany

Key Words: Bayesian regularisation, hierarchical models, high-dimensional data, non-Gaussian ridge regression, scale mixtures of normals

We present a Bayesian approach to generalized regression situations that combines smoothing of nonparametric functions and spatial effects based on smoothness priors with regularization of high-dimensional covariate effects based on shrinkage priors. A general class of smoothing priors is given by multivariate correlated Gaussian distributions, where the precision matrix determines smoothness based on adjacency information. Shrinkage priors are obtained by considering i.i.d. priors with i.i.d. Gaussian priors corresponding to ridge regression as the simplest example. A hierarchical formulation of (scale) mixtures of normals or mixtures of smoothing variances yields a comprehensive class of priors that includes the LASSO as a special case. Due to the hierarchical formulation, existing MCMC algorithms can be adapted with only slight modifications.

166 Challenges of Statistical Inference in 'Large p, Small n' Problems ●▲

Biometrics Section, Section on Nonparametric Statistics, Biopharmaceutical Section, Section on Statistics in Epidemiology, WNAR, IMS

Monday, August 4, 2:00 p.m.–3:50 p.m.

Large P, Small N Asymptotics for Detecting Gene Interactions

✱Michael Kosorok, The University of North Carolina at Chapel Hill, 3101 McGavran-Greenberg Hall, CB 7420, Chapel Hill, NC 27599-7420, *kosorok@unc.edu*; Shuangge Ma, Yale University

Key Words: empirical processes, large p small n data, microarrays, gene networks, high dimensional data, asymptotics

In this talk, we extend previous work on large p, small n asymptotics for significance analysis of individual genes to detecting interactions between genes in microarray experiments. In this setting, n is the number of microarrays and p, the number of genes, is typically orders of magnitude larger than n. The new results yield computationally fast approximations which are accurate enough to correctly identify gene interactions of scientific interest. The overall goal is to determine biologically meaningful networks of genes. The effectiveness of the new methods for identifying networks of genes is demonstrated both theoretically and numerically and compared to applicable previous approaches.

Large Margin Hierarchical Classification

Huxiang Wang, The University of Minnesota; ✱Xiaotong Shen, The University of Minnesota, School of Statistics, Minneapolis, MN 55455, *xshen@stat.umn.edu*; Wei Pan, The University of Minnesota

In this talk, we will present a large margin method for hierarchical classification. The main focus here is to utilize the inter-class dependency to improve the classification performance of flat classification. In such a situation, flat classification is infeasible in the presence of a large number of classes, which occurs often in gene function discovery. Various hierarchical losses will be discussed, in addition to some numerical examples in gene function prediction.

Sufficient Dimension Reduction for Small N, Large P Regressions

✱Lexin Li, North Carolina State University, Department of Statistics, Box 8203. 220-G Patterson Hall, Raleigh, NC 27695, *li@stat.ncsu.edu*; R. Dennis Cook, The University of Minnesota; Chih-Ling Tsai, University of California, Davis

Key Words: Partial least squares, Single-index model, Sliced inverse regression

In regressions with a vector of quantitative predictors, sufficient dimension reduction methods can effectively reduce the predictor dimension, while preserving full regression information and assuming no parametric model. However, current reduction methods require the sample size n to be greater than the number of predictors p. It is well known that partial least squares can deal with problems with $n < p$. In this talk, we first establish a link between partial least squares and sufficient dimension reduction framework. Motivated by this link, we then propose a new dimension reduction method that works for $n < p$ regressions. Both simulations and real data analysis will be presented to demonstrate effectiveness of the proposed method.

Penalized Model-Based Clustering with Application to Microarray Data

✱Wei Pan, The University of Minnesota, A460 Mayo, MMC 303, Biostatistics, SPH, Minneapolis, MN 55455, *weip@biostat.umn.edu*; Benhuai Xie, The University of Minnesota; Xiaotong Shen, The University of Minnesota

Key Words: BIC, EM algorithm, Mixture model, Penalization, Variable selection, Shrinkage

Clustering analysis is one of the most widely used statistical tools in many emerging areas such as microarray data analysis. For microarray and other high-dimensional data, the presence of many noise variables may mask underlying clustering structures. Hence removing noise variables via variable selection is necessary. For simultaneous variable selection and parameter estimation, penalized model-based clustering has been proposed. However, existing penalized methods depend on the assumption of a common diagonal covariance matrix. We aim to relax this assumption and propose a new method. Numerical examples, including application to a microarray data, are provided to demonstrate the utility of the proposed method.

167 Understanding and Improving Communication with Survey Respondents ●▲

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Monday, August 4, 2:00 p.m.–3:50 p.m.

Keeping Up with Survey Respondents

✱Stanley R. Freedman, Energy Information Administration, 1000 Independence Avenue SW, Washington, DC 20585, *stanley.freedman@eia.doe.gov*

Key Words: Survey Methodolgy, Respondent Contacts, Site Visits

The Energy Information Administration (EIA) has taken several steps to improve its communication with survey respondents. In the past 30 years the energy industry has changed substantially because of deregulation and the restructuring of industry business practices. Even though EIA's business surveys are mandatory, we have had to change our approach to communicating with respondents to maintain our response rates and a high level of data quality. Our efforts have focused largely on pre-survey notifications of survey changes, providing multiple modes of data collection, making our surveys correspond to business record keeping practices, and insuring the confidentiality of sensitive price information. Since many of our largest respondents complete multiple EIA surveys, this paper will also highlight EIA's move toward a common look and feel for its internet data collection systems.

Communicating with Survey Respondents at the UK Office for National Statistics

✱Jacqui Jones, U.K. Office for National Statistics, Room 1.126, Government Buildings, Cardiff Road, Newport, NP10 United Kingdom, *jacqui.jones@ons.gov.uk*; Ann Lewis, UK Office for National Statistics; Georgina Jones, UK Office for National Statistics; Jackie Byard, UK Office for National Statistics

Key Words: respondent communication

The mission of the UK Office for National Statistics (ONS) is to improve understanding of life in the United Kingdom and enable informed decisions through trusted, relevant and independent statistics and analysis. The ONS is therefore responsible for conducting, analyzing, and disseminating a broad range of social surveys, business surveys, and the decennial census. Respondent communication is regarded as one of the central elements to facilitating participation in our surveys.This paper will review, compare, and discuss social and business survey respondent communication at ONS. The paper will include an overview of the types, modes, visual design, and content themes of our respondent communication.

Customer and Respondent Outreach Initiatives at the Bureau of Labor Statistics

✱William J. Wiatrowski, Bureau of Labor Statistics, 2 Massachusetts Ave NE room 4130, Washington, DC 20212, *wiatrowski_w@bls.gov*

Key Words: customers, respondents, outreach

The Bureau of Labor Statistics (BLS) provides key economic statistics such as payroll employment, the unemployment rate, and the consumer price index, as well as statistics on compensation, productivity, workplace safety and health, and other topics. The BLS has undertaken a number of customer relations initiatives in recent years designed both to inform data users and to encourage respondents to provide data. Among the topics to be discussed are new publications, Internet sites geared toward respondents, improvements in electronic data collection, and a test to determine what effect targeted changes to materials provided to respondents had on survey response rates.

168 Statistical Issues for Internet Marketing and Price Optimization ●▲

Section on Statistics and Marketing

Monday, August 4, 2:00 p.m.–3:50 p.m.

Internet Advertising and Monetization

✱Paul Dagum, Business.Com, 2120 Colorado Avenue, 3rd Floor, Santa Monica, CA 90404, *pdagum@business.com*

Key Words: search marketing, internet advertising, auction, matching algorithms, monetization

Internet advertising has grown into a multibillion dollar industry. Much of this rapid growth is attributable to search marketing and the creation of virtual marketplaces for long-tail products and services. In search marketing, the search engine plays the role of the market maker matching queries, or supply, to advertisers, or demand. Types of matching algorithms that have been used included first price auctions and generalized second price auctions. Estimation of ad click-through-rate by position and by query type is central to how efficient the search engine can monetize a marketplace and to the equilibrium prices of the auction. We present our experience with our ad auction platform at Business.com, the leading business-to-business search engine. We discuss the challenges of inferring click-through-rates with sparse data and their effect on market prices and match efficiency.

Online Ad Auctions

✱Hal Varian, Google, Inc., 1600 Amphitheatre Parkway, Mt View, CA 94043, *hal@google.com*

Key Words: search engines, online advertising, game theory, auction theory

Several commercial search engines use an auction to sell ad space in which the highest bidders get the most preferred positions. Game theoretic models can be used to predict the equilibrium outcomes for these auctions. It appears that data seem to fit the predictions of the game theoretic models reasonably well.

Modeling Consumer Search for Making Online Advertising Decisions

✱Alan Montgomery, Carnegie Mellon University, Tepper School of Business, 5000 Forbes Ave., Pittsburgh, PA 15241, *alanmontgomery@cmu.edu*

Key Words: Clickstream, Bayesian Models, Purchase Conversion, Text Modeling, Marketing

Many consumers begin their purchase process at search engines such as Google, Yahoo, or MSN instead of traditional retailers. Consumers rely upon the search results provided by these engines along with paid advertising to make decisions about what sites to visit and subsequently which products to purchase. In this study we propose a statistical model that predicts consumer search and the probability of purchase using clickstream data collected from an online sample of consumers. A challenge in analyzing this data is the textual nature of the search strings and the scarcity of many search terms. We also consider how consumers will search based upon the specificity of the search term. We illustrate how this model can be used to aid advertisers in making decisions about how much to bid, what phrase to bid upon, and the appropriate landing page for the consumer once they enter the web site.

169 Research in Health-Related Monitoring ●

Section on Quality and Productivity
Monday, August 4, 2:00 p.m.–3:50 p.m.

Research Issues and Ideas on Health-Related Surveillance

✱William H. Woodall, Virginia Polytechnic Institute and State University, Department of Statistics, Virginia Tech, Blacksburg, VA 24061-0439, *bwoodall@vt.edu*

Key Words: control charts, CUSUM chart, public health surveillance, risk adjustment, statistical process control, spatiotemporal monitoring

In this overview paper, some of the surveillance methods and metrics used in health-related applications are described and contrasted with those used in industrial practice. Many of these health-related surveillance methods are based on the concepts and methods of statistical process control. Public health data often include spatial information, as well as temporal information, so the public health applications can be even more challenging than industrial applications. Research ideas are suggested on various topics in health-related monitoring.

Directionally Sensitive Multivariate Control Charts with an Application to Biosurveillance

✱Inbal Yahav, University of Maryland, R.H. Smith School of Business, College Park, MD 20742, *iyahav@rhsmith.umd.edu*; Galit Shmueli, University of Maryland

Key Words: Multivariate, Control Charts, Hotelling, MEWMA, Performance Analysis, directional sensitivity

Multivariate control charts are used for monitoring multiple series simultaneously. Traditionally, they are aimed at detecting shifts in the mean in any direction. However, in the context of disease outbreak detection, the interest is in detecting an increase in the process means only. Several approaches were suggested for deriving one-directional multivariate Shewhart (Hotelling) charts, but there has not been an extensive comparison of these methods' performance. We compare two approaches: Follmann's correction and Testik's and Runger's (TR) quadratic programming. We also derive directionally sensitive MEWMA charts. We perform an extensive performance analysis of the different methods and their robustness to assumptions. Our results show that TR's approach performs slightly better for normally distributed data, yet Follmann's is more robust to normality and independence assumptions.

Risk-Adjusted Monitoring of Survival Times

✱Landon Sego, Pacific Northwest National Laboratory, P.O. Box 999, MS K6-08, Richland, WA 99352, *landon.sego@pnl.gov*; Marion R. Reynolds, Jr., Virginia Polytechnic Institute and State University; William H. Woodall, Virginia Polytechnic Institute and State University

Key Words: control chart, CUSUM, monitoring, risk adjustment, survival times, censoring

We consider the monitoring of surgical outcomes, where each patient has a different risk of post-operative mortality due to risk factors which exist prior to the surgery. We propose a risk-adjusted survival time CUSUM chart (RAST CUSUM) for monitoring a continuous, time-to-event variable that may be right censored. We compare the ARL performance of the RAST CUSUM chart to the risk-adjusted Bernoulli CUSUM chart, using data from cardiac surgeries to motivate the details of the comparison. The comparisons show that the RAST CUSUM chart is more efficient at detecting a sudden increase in the odds of mortality than the risk-adjusted Bernoulli CUSUM chart, especially when the fraction of censored observations is relatively low or when a small increase in the odds of mortality occurs.

Sparse Time Series Monitoring for Biosurveillance

✱Howard S. Burkom, Johns Hopkins University, Applied Physics Laboratory, Mailstop 8-220, 11100 Johns Hopkins Road, Laurel, MD 20723, *Howard.Burkom@jhuapl.edu*; Yevgeniy Elbert, Johns Hopkins University

Key Words: Sparse Time Series, Biosurveillance, Outbreak Detection

This presentation discusses the problem of detecting small-scale events in biosurveillance data that are relatively sparse in the sense that the median count of monitored time series values is zero. Sparseness may result from syndromic classification that filters out most data records and/or fine spatial resolution into multiple time series. Research goals are to specify scenarios when methods adapted for sparseness are warranted, to examine adaptations of control charts and other algorithms under these scenarios, and to compare the detection performance of these algorithms. Methods considered include CUSUM, EWMA, and temporal scan statistics adaptations. Recent articles have reached differing conclusions in these comparisons in disparate applications. We will apply the chosen methods to authentic and simulated data streams to determine practical guidance for public health monitors.

170 Home Field Advantage in Sports ●▲

Section on Statistics in Sports, *CHANCE*
Monday, August 4, 2:00 p.m.–3:50 p.m.

There's No Place Like Home: Estimating Intra-Conference Home Field Advantage in College Football Using a Longitudinal Model

✱Byron Gajewski, Kansas University Medical Center, School of Nursing, 3901 Rainbow Boulevard, Kansas City, KS 66160, *bgajewski@kumc.edu*

Key Words: longitudinal model, home field advantage

This talk presents a method to measure the impact of the home field advantage for intra-conference college football. The method models longitudinal data across several years while utilizing a unique home field parameter for each individual team. Additionally, two novel yet intuitive measures of home field advantage are proposed. As a case study of the method and the definitions of home field advantage, teams with the best and worst home field advantages within their respective conferences are determined. I will use the model to explore: How does home field advantage change over time? Can we model home field advantage as a multiple of team strength? What is the impact when stadiums are remodeled?

Using Neutral Site Games To Understand Home Field Advantage: Is It Possible?

Gary Simon, New York University; ✱Jeffrey Simonoff, New York University, New York, NY 10012, *jsimonof@stern.nyu.edu*

Key Words: home field advantage, 'last licks', baseball, travel effects, softball

Home field advantage, the finding that home teams win more than half of the games played in a balanced home/away schedules, is well documented in many sports. Four factors are typically cited to account for this: crowd factors, familiarity with local conditions, travel factors, and rule differences for the home versus the visiting team. The simultaneous occurrence of these factors in games played at a team's home field makes it a challenge to sort out each factor's contribution to home field advantage, but neutral site games (where neither team is playing at home) helps address this problem, since crowd factors and local familiarity are much less relevant. Using neutral site NCAA baseball and softball playoff results, we find no evidence that batting last in a baseball or softball game provides an advantage, and find little evidence for the existence of travel fatigue effects.

Coors Field: Why Is the Home Field Advantage so High and What Are Its Implications?

✱Michael J. Schell, Moffitt Cancer Center, 12902 Magnolia Drive, Tampa, FL 33612-9416, *michael.schell@moffitt.org*; Dan Ayers, Vanderbilt University

Key Words: home field advantage, baseball

Coors Field has the greatest home field advantage differentials of any baseball team in major league history. Home field differentials, however, need to be evaluated on the overall quality of the team. A key question is: does the team play particularly well at home, poorly on the road, or both? Data from the 13 seasons at Coors Field suggests that the team bats particularly poorly on the road. Home batting performance can be divided into two periods: pre- and post-humidor—the baseball humidifier introduced in 2002. Pre-humidor averages dropped 30 points after the 7th home game, but post-humidor homestands been more stable. It is possible that the high altitude conditions at Coors disrupts the batting styles of hitters and leads to the team's poor road performance. These issues complicate the Rockies playoff chances, making last year's NL pennant win noteworthy.

171 Analysis of Massive Online Social Networks ●

Section on Statistical Computing, Section on Statisticians in Defense and National Security, Social Statistics Section

Monday, August 4, 2:00 p.m.–3:50 p.m.

Cost-Effective Outbreak Detection in Networks

✱Jure Leskovec, Carnegie Mellon University, *jure@cs.cmu.edu*

Key Words: Social networks, Information cascades, Virus propagation, Sensor Placement, Submodular Functions

In a water distribution network, where should we place sensors to quickly detect contaminants? Or, which blogs should we read to avoid missing important stories? Outbreak detection in both problems can be modeled as selecting nodes in a network in order to detect the spreading of a virus or information as quickly as possible. We present a general methodology for near optimal sensor placement in these and related problems. We exploit the property of submodularity to develop an efficient algorithm that scales to large problems, achieving near optimal placements, while being 700 times faster than a simple greedy algorithm. We evaluate our approach on the two cases mentioned above, showing how the approach leads to deeper insights in both applications, answering criteria trade-off, cost-sensitivity and generalization questions.

Data Analysis at Facebook

✱Jeff Hammerbacher, Facebook, *jeff.hammerbacher@gmail.com*

Key Words: online social networks, sentiment analysis, massive data algorithms

Every day, several terabytes of data are generated by users of Facebook. To extract meaningful information from this data requires a novel infrastructure, scalable algorithms, and respect for the privacy of our users. We will discuss some functional areas in which techniques from statistics and machine learning are applied to Facebook data: user growth and activity, site integrity, influence and sentiment analysis, news feed quality, and platform application health and virility. The focus will be on methodology, rather than results.

Deriving Marketing Intelligence from Online Discussion

✱Natalie Glance, Google - Pittsburgh, 4720 Forbes Avenue, Pittsburgh, PA 15213, *nglance@gmail.com*

Weblogs and message boards provide online forums for discussion that record the voice of the public. Woven into this mass of discussion is a wide range of opinion about consumer products, movies and other topics. This presents an opportunity for companies to understand and respond to the consumer by analyzing this unsolicited feedback. Given the volume, format and content of the data, the appropriate approach to understand this data is to use large-scale web and text data mining technologies.

172 Applications of Copulas in Biostatistics and Statistical Genetics ●

ENAR, Biopharmaceutical Section, WNAR, SSC, Biometrics Section

Monday, August 4, 2:00 p.m.–3:50 p.m.

A General Introduction to Frailty Models for Multivariate Survival Models and Induced Copula Structures

✱David Oakes, University of Rochester, Dept of Biostatistics and Computational Biology, 601 Elmwood Avenue Box 630, Rochester, NY 14642, *oakes@bst.rochester.edu*

Key Words: archimedean copulas, censoring, Clayton's model, Kendall's tau, truncation

In survival analysis a frailty is an unobserved random effect. Usually this is taken to act multiplicatively on the hazard function. Introduction of frailties into univariate survival models typically has the effect of attenuating effects of fixed covariates. In multivariate survival analysis, introduction of a common frailty, or of correlated frailties, among related individuals is a useful approach to modeling dependence between their survival times. Features of the frailty distribution are reflected in dependence characteristics of the observable joint survival distribution. There is a very natural connection with copula models and with nonparametric measures of dependence. The talk will discuss these connections and certain properties, such as invariance under censoring and truncation, that are important in biostatistical applications of copula models.

Modeling Familial Association of Ages at Onset in the Presence of Competing Risk

✱Joanna Shih, National Cancer Institute, Bethesda, MD 20892-7434, *jshih@mail.nih.gov*

Key Words: Familial association, Competing risk, Age at onset, Cross-ratio, Odds-ratio

In genetic family studies, ages at onset of diseases are routinely collected. Often one is interested in assessing the familial association of ages at onset of a certain disease type. However, when the competing risk is present and related to the disease of interest, the usual measure of association by treating the competing event as an independent censoring event is biased. We propose a bivariate model that incorporates two types of association: one is between the first event times of paired members, and the other is between the failure types. We consider flexible measures for both types of association, and propose an estimate procedure for these measures. The proposed method is illustrated using the kinship data from the Washington Ashkenazi Study.

Quantitative Trait Linkage and Association Analysis Using Gaussian Copulas

✱Mingyao Li, University of Pennsylvania, 624 Blockley Hall, Department of Biostatistics and Epidemiology, Philadelphia, PA 19104, *mingyao@mail.med.upenn.edu*; Michael Boehnke, The University of Michigan; Goncalo Abecasis, The University of Michigan; Peter Song, University of Waterloo

Key Words: quantitative traits, linkage, association, copula, statistical genetics

Mapping variants that influence quantitative traits is an important problem in genetics studies. Traditional QTL mapping relies on a variance-components (VC) approach with the assumption that trait values in a family follow a multivariate normal distribution. Violation of this assumption can lead to inflated type I error, reduced power, and biased parameter estimates. To accommodate nonnormally distributed data, we describe a "copula VC method" that directly models the nonnormal distribution using Gaussian copulas. Our method allows the analysis of continuous, discrete, and censored data, and the standard VC method is a special case when the data are distributed as multivariate normal. Through the use of link functions in generalized linear models, the copula VC method can easily incorporate covariates, and therefore provides a unified framework for linkage and association analysis.

173 Are Disability Statistics Relevant and Useful for National Policies and Programs? ●▲

Committee on Statistics and Disability, Section on Government Statistics, Social Statistics Section, Scientific and Public Affairs Advisory Committee
Monday, August 4, 2:00 p.m.–3:50 p.m.

Are Disability Statistics Relevant and Useful for National Policies and Programs?

✱Siobhan Carey, Central Statistics Office, Social and Demographic Statistics, Cork, Ireland, *siobhan.carey@cso.ie*; ✱Michele J. Connolly, Sweetgrass Consulting LLC, 9556 Highwind Court, Columbia, MD 21045, *michelebabb@verizon.net*; ✱Daniel Mont, World Bank, Disability and Development Team, 1818 H Street N.W., Washington, DC 20433, *dmont@worldbank.org*; ✱Susan Stobert, Statistics Canada, Participation and Activities Limitations Survey, Jean Talon Bldg, Tuney's Pasture, Ottawa, ON K1A0T6 Canada, *susan.stobert@statcan.ca*

Key Words: human functioning, disability, statistical methods, frameworks, official statistics, development planning

This session asks whether national statistical offices have yet achieved a viable framework for the successful development of disability statistics necessary for a broad range of national planning, policy and public education purposes. It introduces the work of national statistical offices concerning the measurement and reporting of human functioning as well as reporting on issues facing nations and their citizens with disabilities. The session highlights the comparative experience of two national statistical offices and one international organization for their effectiveness in preparing valid and reliable official statistics useful for development planning.

174 To the Nth Power: Younger Statisticians Taking the Lead ●▲

Social Statistics Section, Section on Government Statistics
Monday, August 4, 2:00 p.m.–3:50 p.m.

To the Nth Power: Younger Statisticians Taking the Lead

✱Juanita Tamayo Lott, U.S. Census Bureau, Human Resources Division, 20233, *juanita.t.lott@census.gov*; ✱Nagambal Shah, Spelman College, 30314-4399, *nshah@spelman.edu*; ✱Ben Hansen, The University of Michigan, Ann Arbor, MI 48109-1092, *ben.hansen@umich.edu*; ✱Gayla Olbricht, Purdue University, 47907-2067, *ghobbs@stat.purdue.edu*; ✱Brian A. Millen, Eli Lilly and Company, 11817 Lancaster Circle, Carmel, IN 46033, *Millen_Brian_A@Lilly.com*

Key Words: non traditional statisticians, leadership initiatives, StatFest, younger statisticians, STATCOM

This session will highlight the changing 'face' of the statistics profession with younger and non-traditional statisticians taking the lead on important initiatives. Panelists will discuss successful strategies for attracting nontraditional statisticians to the profession and increasing the visibility of the field via StatFest, Statistics in the Community (STATCOM), and other emerging programs and networks created or led by younger statisticians. Panelists will also address effective and innovative strategies to ensure that this pool develops and moves into mid-level and senior-level positions.

175 Pedagogical Issues in an Introductory Statistics Course

Section on Statistical Education
Monday, August 4, 2:00 p.m.–3:50 p.m.

Pedagogical Issues in Introductory Statistics

✱Carolyn Cuff, Westminster College, Hoyt Science Center, New Wilmington, PA 16172-0001, *ccuff@westminster.edu*; ✱E. Jacquelin Dietz, Meredith College, 3800 Hillsborough Street, Raleigh, NC 27607-5298, *dietzjac@meredith.edu*; ✱Kim Gilbert, Georgia Gwinnett College, 1000 University Center Lane, Lawrenceville, GA 30043, *kgilbert@ggc.usg.edu*; ✱Robert Gould, University of California, Los Angeles, Department of Statistics, Los Angeles, CA 90095-1554, *rgould@stat.ucla.edu*; ✱Thomas H. Short, Indiana University of Pennsylvania, Mathematics Department, 210 South 10th Street, Indiana, PA 15705, *tshort@iup.edu*

Key Words: Pedagogy, Technology, Education, Introductory Statistics, Assessment

A panel of experts shares ideas and experiences to improve pedagogy in the first statistics course. Topics will include teaching with activities, using assessment to improve learning, incorporating technology into the course, and which topics to include.

176 New Statistical Methods and Challenges in Neurodegenerative Disease Research ●

Section on Statistics in Epidemiology, WNAR, Biometrics Section

Monday, August 4, 2:00 p.m.–3:50 p.m.

Evaluating the Accuracy of a Diagnostic Test When the Diagnostic Test Is Subject to Measurement Error Using an Internal Reliability Sample, with Applications to Alzheimer's Disease

✱Sharon X. Xie, University of Pennsylvania, 423 Guardian Drive, Philadelphia, PA 19104, *sxie@mail.med.upenn.edu*; Christopher M. Clark, University of Pennsylvania

Key Words: Sensitivity, Specificity, ROC curve, Measurement error, Alzheimer's disease

Cerebrospinal fluid (CSF) tau is useful in detecting Alzheimer's disease (AD). It is often assumed to measure the true biological quantity without error when its sensitivity and specificity are calculated. However, in practice this assumption almost never holds. When an internal reliability sample is available for CSF tau, we have shown that averaging replicates and ignoring measurement error can sometimes lead to serious biases of sensitivity and specificity estimates. We present a bias-correction approach to remove the biases of sensitivity and specificity estimates introduced by measurement error based on an internal reliability sample. Asymptotic distributions were obtained for the proposed estimators. Extensive simulations were conducted to evaluate the proposed approach. All methods are illustrated using the University of Pennsylvania AD biomarker study.

Biologically Driven Data Reduction of Patterns of Change Seen on Structural Imaging

✱Danielle Harvey, University of California, Davis, 1442 Gravink Ct, Woodland, CA 95776, *djharvey@ucdavis.edu*; Laurel A. Beckett, University of California, Davis

Key Words: spatial decomposition, MRI data, longitudinal data

Alzheimer's disease affects the underlying structure and tissue of the brain. Magnetic Resonance Imaging enables researchers to observe tissue atrophy and abnormalities in the brain. These images may be broken up into over 200,000 tiny volumes, or voxels, of information. Larger studies are now focusing on serial images to detect disease-related patterns of brain changes. Typical analyses have reduced these data to a one-number summary, such as a volume, repeated over time. However, such a summary may not adequately capture the ways in which the brain changes. Alternative summaries may include decompositions of the serial voxel data that identify key features of change in size, location, and shape. We present a possible strategy for capturing patterns of change along with simulation results to illustrate the performance of such a strategy.

Comparison of Biomarkers for Alzheimer's Disease Progression: Insights from the Alzheimer's Disease Neuroimaging Initiative

Laurel A. Beckett, University of California, Davis; ✱Hao Zhang, University of California, Davis, Department of Statistics, Davis, CA 95616, *hzhang@wald.ucdavis.edu*

Key Words: Alzheimer's Disease, ADNI, Biomarker, Longitudinal, Neuroimaging

The aim of the Alzheimer's Disease Neuroimaging Initiative (ADNI) is to evaluate potential biomarkers for clinical disease progression, using neuroimaging, serum and cerebrospinal fluid testing. The ultimate goal is to have more precise assessment not just of the natural history of Alzheimer's disease (AD), but also of the impact of interventions, so that we can speed the development, testing and delivery of treatments. Comparison of potential biomarkers is challenging in part because the marker data are correlated longitudinal sequences with potentially informative missing values. Moreover, we would like to compare across multiple clinical outcomes: conversion from normal to mild cognitive impairment or mild cognitive impairment to AD; longitudinal change in multiple neuropsychological or functional assessment measures.

Models for Identifying Risk Factors for Transitions from Intact Cognition to Mild Cognitive Impairment and Dementia

✱Richard Kryscio, University of Kentucky, 230 Sanders-Brown Building, 800 S. Limestone St., Lexington, KY 40536, *kryscio@email.uky.edu*; Lei Yu, University of Kentucky; Suzanne Tyas, University of Waterloo; David Snowdon, University of Kentucky

Key Words: Alzheimer's disease, mild cognitive impairment, transition models, risk factors

We review strategies for modeling panel data collected on the cognitive status of elderly subjects being followed for the occurrence of dementia. Response at any given wave is a nominal variable representing the states cognitively intact, mild cognitive impairment or a deficit in activities of daily living. Interest centers about transitions among these states with eventual absorption into one of two competing states: dementia or death. We compare the relative merits of several modeling strategies that will allow estimation of the effect of covariates on these transitions This includes Markov transition models, and competing risk models. An application to the NUN study data, a longitudinal cohort of 678 sisters of Notre Dame aged 75 or older at baseline, will illustrate the results.

177 Adaptive Design and Dose-Finding in Clinical Trials ●▲

Biopharmaceutical Section, Section on Risk Analysis, WNAR, Biometrics Section
Monday, August 4, 2:00 p.m.–3:50 p.m.

Normal Dynamic Linear Model for Dose-Finding in a Phase II Bayesian Adaptive Trial

✱Jason T. Connor, Berry Consultants, 2534 Lake Debra Drive, #108, Orlando, FL 32835, *jason@berryconsultants.com*; Scott M. Berry, Berry Consultants; Donald A. Berry, M.D. Anderson Cancer Center

Key Words: Adaptive trial, NDLM, Bayesian, Biostatistics, Clinical trial

We describe a normal dynamic model for dosing finding. This Gaussian process model is a flexible not necessarily monotonic model for outcomes primarily defined by one parameter (e.g., mean of a continuous response, Bernoulli probability, mean survival). Using the NDLM we estimate the probability distribution of the dose response curve and adaptively randomize to efficiently identify the dose or doses of interest (minimally effective dose, ED90, maximum utility dose, etc). The dose-response curve has no forced parametric form however there is sharing across doses governed by the drift parameter. We compare the NDLM to common parametric dose-response curves and demonstrate the flexibility and strength of the NDLM model.

Adaptive Dose Selection Using Interim Analyses When Fitting a Bayesian Emax Model to Clinical Trial Data

✱Neal Thomas, Pfizer, Inc., 50 Pequot Avenue, New London, CT 06320, *neal.thomas@pfizer.com*; Byron Jones, Pfizer, Inc.; Patrick Johnson, Pfizer, Inc.; Helen Richardson, Pfizer, Inc.

Key Words: Emax model, dose adaptive design, dose ranging clinical trial, Bayes, stopping rules

The presentation will describe an adaptive dose response study to compare the mean effects of a set of doses with placebo using a nonlinear (Emax) function. The design has two parts: the first part is a small parallel group trial followed by a second part which is adaptive in nature. The initial selection of doses is made using optimal design theory with power to detect a difference of a specified size between the highest dose and placebo. Subsequent patients are recruited in cohorts and the data are reanalysed using a Bayesian nonlinear model. The doses for each cohort are selected using Bayesian posterior probabilities of satisfying various criteria for success. The Bayesian posterior probabilities will also be used to stop the trial early for success or futility if necessary. The operating characteristics of the adaptive design will be assessed.

Response-Adaptive Dose Finding Combing Multiple Comparison and Modeling Approaches

✱Frank Bretz, Novartis Pharmaceuticals, *frank.bretz@novartis.com*; Jose Pinheiro, Novartis Pharmaceuticals; Bjoern Bornkamp, University of Dortmund

Key Words: dose-response, dose-finding, contrast, multiple testing, nonlinear regression, adaptation

Adaptive dose-ranging studies form an innovative class of designs aimed at striking the desired balance between needed dose response information and increased costs and timelines. In these designs, the number of doses and/or proportion of patients allocated to doses are allowed to change as safety and efficacy information is accrued and analyzed. This presentation will describe an adaptive dose-ranging approach based on a hybrid methodology combining multiple comparison procedures (MCP) and modeling, the so-called MCP-Mod method. In the response-adaptive version of MCP-Mod, dose allocation ratios are recalculated at each interim analysis, based on the accumulated information, according to a Bayesian optimality criterion which accounts for model uncertainty. The methodology will be compared to alternative dose ranging strategies via simulation.

Comparing a Bayesian Approach with the Frequentist T-Statistic Method in Adaptive Dose-Finding Trials

✱Nitin R. Patel, Cytel Inc., 675 Massachusetts Avenue, Cambridge, MA 02139, *nitin@cytel.com*; James Bolognese, Merck Research Laboratories; Inna Perevozskaya, Merck Research Laboratories; Robert B. Smith, Cytel Inc.

Key Words: adaptive, dose finding, Bayesian, t-statistic method, frequentist, simulation

There are several adaptive designs for dose-finding trials that modify randomization ratios in cohorts based on accumulating information of responses to improve on standard fixed ratio designs. We will describe results from simulation experiments that compare two adaptive methods: a Bayesian method that uses a variance reduction criterion to adaptively allocate doses developed by Scott Berry, and the t-statistic method of Ivanova, Bolognese, and Perevozskaya. The methods will be compared for data generated from a four-parameter logistic mean response curve for both continuous and binary endpoints. Effectiveness along dimensions such as power, bias, dose-response estimation and probability of finding the target dose will be examined. Two definitions of the target dose will be investigated: the dose that gives a specified magnitude of response, or a specified difference from placebo.

178 Bayesian Bioinformatics and Bayesian Biostatistics ●

Section on Bayesian Statistical Science, Biopharmaceutical Section, WNAR
Monday, August 4, 2:00 p.m.–3:50 p.m.

Combining Classifiers To Select Differentially Expressed Genes for Microarray Data

✱Lynn Kuo, University of Connecticut, 215 Glenbrook Rd, Storrs, CT 06269, *lynn@stat.uconn.edu*; Wangang Xie, University of Connecticut; Yifang Zhao, University of Connecticut

Key Words: SAM, LIMMA, MAANOVA, semiparametric hierarchical Bayesian, EM, MCMC

Many statistical tests are being used to select differentially expressed genes for microarray data analysis. It would be useful to evaluate the sensitivity and specificity of each test given the test results. More importantly, it would be desirable to combine these results intelligently to achieve better power. We propose a model with likelihood based inference that would allow us to achieve both goals in an integrated and comprehensive fashion. Both EM and Bayesian methods are developed for it.

Detection of Copy Number Variations from High-Density SNP Arrays: An Integrated Bayesian Hidden Markov Model Approach Incorporating Pedigree Information

✱Zhen Chen, University of Pennsylvania, Philadelphia, PA , *chenz@mail.med.upenn.edu*; Mahlet Tadesse, Georgetown University; Kai Wang, University of Pennsylvania; Mingyao Li, University of Pennsylvania

Key Words: Hidden markov model, MCMC, Copy number variations, Bayesian, Genome, Model uncertainty

Copy number variations (CNVs) refer to gains and losses of genomic elements compared to a reference genome assembly. Studies have demonstrated the heritability of CNVs but few incorporated family structures. We develop an integrated Bayesian approach that aims to incorporate family relationships when inferring CNVs. We assume copy number sequence along the chromosome follows a Markov model with transition probabilities dependent on genetic distances between adjacent SNPs. We also allow for de novo events in offspring's CNV calls and use another HMM to account for the dependence with neighboring SNPs in the same de novo CNV region. Our approach yields posterior distributions of CNV configurations, thus providing an uncertainty measure for the inferred CNVs. We evaluate the performance of the method by applying it to simulated datasets and the CEU trio data from HapMap.

Reexamine Bayes Factor Methods for Bayesian Phylogenetic Inference

Ming-Hui Chen, University of Connecticut; Yu Fan, University of Connecticut; Paul O. Lewis, University of Connecticut; Lynn Kuo, University of Connecticut; ✱Wangang Xie, University of Connecticut, 06269, *wxie@stat.uconn.edu*

Key Words: Bayesian phylogenetics, Bayes factor, Marginal likelihood, Improved monte carlo methods

Several methods have been advocated to estimate the marginal likelihoods in the Bayes factor computation to infer phylogenies. Brown and Lemmon (2007) advocated the harmonic mean method because of its desirable statistical behavior from their empirically based simulations. Lartillot and Philippe (2006) advocated thermodynamic integration method because of its accuracy. In this paper, we reexamine this controversy. Moreover, we will consider other methods for computing Bayes factors. They include the Laplace approximation, path importance sampling, path bridge sampling (a combination of bridge and path sampling), and warp bridge sampling. Results from both simulation and real data analysis will be reported.

Incorporating Structural Equation Modeling to Fecundity Models

✱Sung Duk Kim, National Institute of Child Health and Human Development, 6100 Executive Boulevard, Rockville, MD 20852, *kims2@mail.nih.gov*; Rajeshwari Sundaram, National Institutes of Health; Germaine B. Louis, National Institute of Child Health and Human Development

Key Words: Conception, Fertility, Latent variables, Markov chain Monte Carlo, Menstrual Cycle, Structural Equations Model

In epidemiologic studies, human fecundability is measured by probability of conception in a menstrual cycle by a couple. We propose a general model for joint modeling of intercourse behavior and human fecundability through a classic conception probability model and structural equation model which involves a set of latent variables to capture dependence between intercourse behavior on consecutive days in a menstrual cycle, so that the proposed model can accommodate not only a broad variety of intercourse patterns and dependency structure, but also general covariate effects, heterogeneity among fecund couples in menstrual cycle viability and in frequency of intercourse. Markov Chain Monte Carlo method is used to carry out Bayesian posterior computation. A detailed analysis of the New York State Angler pregnancy data is presented to illustrate the proposed methodology.

Statistical Methods for Analysis of Genomic Data with Graphical Structures

✱Caiyan Li, University of Pennsylvania, Department of Biostatistics & Epidemiology, Philadelphia, PA 19104-6021, *licaiyan@mail.med.upenn.edu*; Hongzhe Li, University of Pennsylvania

Key Words: Gene expression, Network, eQTL, Markov random field, Regularization

Graphs and networks are common ways of depicting information. In biology in particular, many different biological processes, such as regulatory networks, metabolic pathways, and protein-protein interaction networks, are represented by graphs. This kind of a priori information gathered over many years of biomedical research is a useful supplement to the standard numerical genomic data such as microarray gene expression data. However, how to efficiently incorporate information encoded by the known biological networks represented as graphs into analysis of various types of numerical genomic data raises interesting statistical and computational challenges. We present several new statistical methods for incorporating network information in to analysis of genomic data, including a network-constrained regularization procedure and a hidden Markov random field approach for analysis of eQTL data.

179 Section on Health Policy Statistics Student Paper Award Winners

Section on Health Policy Statistics, Section on Nonparametric Statistics

Monday, August 4, 2:00 p.m.–3:50 p.m.

An Algorithm for Optimal Tapered Matching, with Application to Disparities in Survival

✱Shoshana R. Daniel, University of Pennsylvania, *skrieger@mail.med.upenn.edu*; Katrina Armstrong, University of Pennsylvania; Jeffrey H. Silber, University of Pennsylvania; Paul R. Rosenbaum, University of Pennsylvania

Key Words: assignment algorithm, combinatorial optimization, matched sampling, tapered matching

In a tapered matched comparison, one group of individuals, called the focal group, is compared to two or more nonoverlapping matched comparison groups constructed from one population in such a way that successive comparison groups increasingly resemble the focal group. An optimally tapered matching solves two problems simultaneously: it optimally divides the single comparison population into nonoverlapping comparison groups and optimally pairs members of the focal group with members of each comparison group. We show how to use the optimal assignment algorithm in a new way to solve the optimally tapered matching problem, with implementation in R. This issue often arises in studies of groups defined by race, gender, or other categorizations such that equitable public policy might require an understanding of the mechanisms that produce disparate outcomes to necessitate reform.

Estimating the Capacity for Improvement in Risk Prediction with a Marker

✱Wen Gu, University of Washington, 98195, *wengu@u.washington.edu*

Key Words: Classification, diagnostic likelihood ratio, posterior probability, logistic regression, diagnostic test, biomarker

Consider a set of baseline predictors X to predict a binary outcome D and let Y be a novel marker or predictor. This paper is concerned with evaluating the performance of the augmented model P(D=1|Y,X) compared with the baseline model P(D=1|X). The diagnostic likelihood ratio, DLR(Y), quantifies the change in risk obtained with knowledge of Y=y for a subject with baseline risk factors X. The notion has been promoted as a way to capture the increment in risk prediction due to Y. We propose methods for making inference about DLRX(Y). We also demonstrate how the population performance of baseline and augmented risk models can be compared using baseline data from a cohort and marker data from a nested case-control study. Finally, we show how the methodology yields estimates of covariate specific predictiveness curves that can be used by an individual to decide the ascertainment of Y.

A Class of Semiparametric Mixture Cure Survival Models with Dependent Censoring

✱Megan Othus, Harvard University/Dana Farber Cancer Institute, 677 Huntington Ave., Building 2, Floor 4, Boston, MA 02115, *mothus@fas.harvard.edu*; Yi Li, Harvard University/Dana Farber Cancer Institute; Ram C. Tiwari, National Cancer Institute

Key Words: transformation model, right censoring, estimating equation

Modern cancer treatments have substantially improved cure rates and have generated a great interest in and need for proper statistical tools to analyze survival data with non-negligible cure fractions. Data with cure fractions are often complicated by dependent censoring, and analysis of this type of data typically involves untestable assumptions about the dependence of the censoring and the survival times. Motivated by the analysis of NCI SEER data, we propose a class of general semiparametric transformation cure models that allows for dependent censoring without making parametric assumptions on the dependence relationship and use the proposed methods to investigate potential racial disparities in prostate cancer cures. The proposed class of models encompasses a number of common models for the latency survival function, including the proportional hazards and proportional odds models.

A New Synthesis Method for Multiple Linear Regression

✱Nan Hu, University of Washington, 846 NE 120th St, Seattle, WA 98125, *nanhu@u.washington.edu*

Key Words: Conditional expectation, Meda-analysis, Linear regression, Prediction models, Synthesis, Delta method

Synthesis method is an important tool to estimate the result in meta analysis in which results of several studies addressing a set of related research hypotheses are combined. Recently, Samsa et al.(2005) proposed a method to combine regression coefficients from univariate linear model into multivariate linear model, and shows the most promise where the primary interest is in generating predicted values for identifying high-risk population. However the validity of this method relies on normality of data and can not provide variance estimate. We proposed a new synthesis method based on conditional expectations of outcome given covariate. The variance estimate of parameter and predicted values was derived using delta method. Simulation study shows that this method performs much better than the method proposed in Samsa et al.(2005). This new method was applied to anlyze blood pressure data.

Semiparametric Efficient Estimator for Incomplete Longitudinal Binary Data with Application to Smoking Trends

✱Jamie Perin, The University of North Carolina at Chapel Hill, *jperin@bios.unc.edu*

Key Words: Dropout, Generalized Estimating Equations, Longitudinal studies, Missing at Random

Generalized estimating equations, while commonly used in the analysis of incomplete longitudinal data, are invalid for data that is not missing completely at random. Robins, Rotnitzky and Zhao (1995) introduced a class of weighted estimating equations that are valid under dropouts missing at random, including an easy-to-implement, but inefficient member. Robins and Rotnitzky (1995) gave the more complex semi-parametric efficient estimator in this class and applied it to continuous data. A specific form of the estimator is developed for binary data and used to assess the efficiency of the simpler estimator in a simulation study. Both are applied in the estimation of cigarette smoking trends in the United States from a cohort of 5,078 young adults. This extends an earlier analysis by Preisser et. al. (2000) who analyzed smoking trends with the inefficient weighted estimating equations.

180 Environmental Risk: Risk Assessment and Dose-Response Models

Section on Risk Analysis, Section on Statistics in Epidemiology

Monday, August 4, 2:00 p.m.–3:50 p.m.

Estimating Daily Growth Estimates of Ride Tide Algae

✱Susan J. Simmons, The University of North Carolina, Wilmington, Department of Mathematics and Statistics, 601 South College Road, Wilmington, NC 28411, *simmonssj@uncw.edu*; Alaina Houmard, University of North Carolina Wilmington

Key Words: Growth curve, Weibull growth curve

For many years, Marine scientists have estimated growth of algae by simply looking at the linear portion of the growth curve. This method produces unreliable results that can dramatically change depending on the portion designated as linear. In this talk, we introduce a more reliable method using estimates from the entire growth curve. The delta method is used to estimate the variance of this new estimator and results from a simulation study will be shown.

Empirical Evaluation of Sufficient Similarity in Dose-Response for Environmental Risk Assessment of a Mixture of 11 Pyrethroids

✱LeAnna G. Stork, Monsanto Company, Mail Zone O3A, 800 N. Lindberegh Blvd., St. Louis, MO 63167, *leanna.g.stork@monsanto.com*; Chris Gennings, Virginia Commonwealth University; Scott Marshall, Virginia Commonwealth University; Mike DeVito, U.S. Environmental Protection Agency; Kevin M. Crofton, U.S. Environmental Protection Agency

Key Words: Equivalence testing, Mixed models, Nonlinear models, Random coefficient

Chemical mixtures in the environment are often the result of a dynamic process. When dose-response data are available on random samples throughout the process, equivalence testing can be used to determine whether the mixtures

are sufficiently similar based on a prespecified biologically important similarity region. Consider the case where dose-response data are available only for a reference mixture from the dynamic process. Following the logic of Stork et al (*JABES*,2008), it is useful to determine how much additional random variation associated with unobserved candidate mixtures from the dynamic process can be added such that the candidate mixtures are sufficiently similar in dose-response to the reference mixture, based on the similarity region. An example is given for a mixture of pyrethroids. This research is not associated with Monsanto. This abstract does not reflect EPA policy.

Model Averaging of Dosages from Toxicology Studies: Approaches and Pitfalls

✱Edward L. Boone, Virginia Commonwealth University, Department of Statistics and Operations Research, 1001 W. Main St., Richmond, VA 23284, *elboone@vcu.edu*

Key Words: Dose Response, Bayesian Methodology, Model Averaging

Model averaging of dosages from toxicology studies have been recently introduced into the literature. When employing these averaging techniques often posterior model probabilities are used which are commonly approximated using BIC, KIC or some other information based criteria. While these approximations are valid asymptotically, there is no information on how accurate these methods are in low to moderate sample size cases. This presentation considers the performance of BIC and KIC against the fully Bayesian approach using posterior model probabilities generated by numerical integration. Examples of how to apply each approach and a comparison of each is considered.

Low-Dose Extrapolation from Points of Departure or from Model-Averaged Estimates

✱A. John Bailer, Miami University, Dept. of Math. & Stat., Oxford, OH 45056, *baileraj@muohio.edu*; Matthew Wheeler, Centers for Disease Control and Prevention

Key Words: risk estimation, benchmark doses

Experiments with relatively high dose conditions are often used to predict risks at appreciably lower doses. A point of departure can be calculated as the dose associated with a moderate response level (e.g. benchmark dose associated with a benchmark response of 10%). A linear extrapolation to lower doses can then follow. This method is compared to the use of model-averaged predictions in the context of a number of animal bioassay experiments.

181 Getting It Real: Encouraging Critical Thinking Through Exposure to Reality ▲

Section on Statistical Education

Monday, August 4, 2:00 p.m.–3:50 p.m.

How To Structure Effective Semester-Long Group Projects for Applying and Integrating Course Concepts

✱Douglas M. Andrews, Wittenberg University, Dept of Math and Comp Sci, Wittenberg University, Springfield, OH 45501, *dandrews@wittenberg.edu*

Key Words: student projects, group projects, critical thinking, statistics education

Long-term student projects can promote the kind of critical thinking required to apply statistical ideas and to integrate the various course components. In particular, students can be involved in all phases of a data-based study—conception, design, implementation, analysis, and reporting. Some instructors balk at student projects, however, often based on initial experiences that result in lame topics, dysfunctional groups, and projects that are either too ambitious or too narrow in scope. The key to avoiding such pitfalls lies in the way that we as instructors structure the administration of the projects and build in opportunities for our intervention and guidance. This talk will include suggestions for 1) multiple, graded, intermediate checkpoints, 2) guidelines and assessment criteria for all phases of the project, and 3) self- and peer-critique of student projects.

Utilizing Projects in Teaching Statistics to Professional Students: Successes and Challenges

✱Heather M. Bush, University of Kentucky, 121 Washington Ave, College of Public Health, 205C, Lexington, KY 40536-0003, *heather.bush@uky.edu*

Key Words: statistics education, projects, data analysis, professional students

Most professional and graduate students take statistics courses to better understand statistics in the literature and in their own research. The use of projects in a second semester biostatistics course for professional and graduate students in fields including medicine, pharmacy, public health, dentistry, and rehabilitation science will be discussed. Emphasis will be on the successes of using individual and group projects to provide students with examples beyond the textbook, a "roadmap" for future analyses, and a better appreciation for the work involved in data analysis. Additionally, handling the challenges of naïve research/programming/analysis skills in students and preventing both the instructor and the students from being overwhelmed will also be addressed.

Student Generated Data vs. Teacher Supplied Real Data vs. Teacher Created Data

✱Steven Dafilou, Springside School, 8000 Cherokee Street, Philadelphia, PA 19118, *sdafilou@springside.org*

This presentation discusses the relative merits of student generated data, teacher supplied real data, and teacher created data. The presentation uses responses from AP Stat and College Intro Stat teachers. It finds that context, pedagogy and timeframe are the key factors in deciding which source to use.

Using Real Data To Teach Statistics

✱Paulette Ceesay, Merck & Co., Inc., 351 N. Sumneytown Pike, North Wales, PA 19454, *paulette_ceesay@merck.com*

Key Words: teaching statistics, data analysis project, critical thinking

The analysis of real life data is a good way to help students appreciate the uses of contemporary applied statistics in various disciplines. The use of a data analysis project as a mechanism to help students develop critical thinking and decision-making ability using quantitative methods and statistical software will be discussed. Emphasis will be on analyzing and interpreting data, as well as, presenting and documenting results. The challenges associated with students generating an appropriate dataset incorporating the statistical methodologies discussed in the course using the internet or other media, personal research or a self-generated survey will be addressed. Data analysis projects can be performed individually or as a group.

Encouraging Statistical Thinking Through the Writing of Newspaper Articles

✱Meike Niederhausen, University of Portland, Department of Mathematics, 5000 N. Willamette Blvd., Portland, OR 97203, *niederha@up.edu*

Key Words: introductory statistics, projects, critical thinking

Group projects are a commonly used tool in statistics classes to encourage students to develop their statistical and critical thinking skills. Students are commonly asked to write a journal-style report at the end of the semester where they are expected to summarize and statistically analyze the results of the data they collected. We have found that students primarily concentrate on the summary of their results and leave very little energy (and time) for the critical analysis. This talk will discuss the benefits of having students write a newspaper-style article rather than a journal-style report. The shorter and more creative article encourages students to move beyond the world of statistical terminology and actually think about what their results mean and how they could be explained to a general audience.

182 Statistical Science Applications in National Security: A Session in Honor of Professor Donald P. Gaver ●▲

Section on Statisticians in Defense and National Security

Monday, August 4, 2:00 p.m.–3:50 p.m.

Donald Gaver's Contributions to Reliability Theory and Applications

✱Francisco J. Samaniego, University of California, Davis, One Shields Avenue, Davis, CA 95616, *fjsamaniego@ucdavis.edu*

Key Words: system reliability, stochastic models, failure time characteristcs, inspection, availability, repair

This presentation will review Don Gaver's many and varied contributions to Reliability. It is filled with realistic engineering scenarios, sophisticated modeling and state-of-the-art statistical tools. The range of Don's interests is uncommonly broad. He has studied a wide variety of systems, the modeling and estimation of their performance, the examination of availability, inspection and repair questions and the application of theoretical findings to the problems that motivated the work. Don's work readily reveals his powerful mathematical capabilities and his strong inclination to work on real problems. Much of his work addresses problems arising in military contexts. His contributions to Reliability span 50 years and have had an enormous impact on defense-related work in this area. It is a pleasure to have the opportunity to review and discuss Don's work. How much time do I have?

Uncertain Time-Critical Tasking Problems

✱Kevin D. Glazebrook, Lancaster University, Department of Management Science, Lancaster, International LA1 4YX United Kingdom, *k.glazebrook@lancaster.ac.uk*

Key Words: Service control, Queueing, Dynamic policies, Dynamic programming

In my long association with Don Gaver, he has consistently shown a healthy disdain for the conventional in model-building and analysis. His commitment is to applicability rather than tractability. In this spirit, and motivated by a plethora of military and other applications (maritime domain awareness), we (and others, including Pat Jacobs) have analyzed classes of service control problems in which conventional assumptions that (i) tasks wait indefinitely for service, and (ii) successful service may be observed instantaneously, are relaxed. The allocation of a large amount of processing to a task may increase the chances of its successful service but may also result in the loss of many unserved tasks from the system. The paper describes simple approaches to the development of dynamic policies for the allocation of service for such systems which may be easily computed.

Heavy Traffic Analysis of Real-Time Queues

✱John Lehoczky, Carnegie Mellon University, Department of Statistics, Pittsburgh, PA 15213, *jpl@stat.cmu.edu*; Kavita Ramanan, Carnegie Mellon University; Steven Shreve, Carnegie Mellon University; Lukasz Kruk, Maria Curie-Sklodowska University

Key Words: queues with deadlines, heavy traffic, Brownian motion, random measures

Don Gaver made many important contributions in the area of queuing theory, especially priority queuing and heavy traffic diffusion approximations. We build on his contributions to present a heavy traffic analysis of real-time queues, queuing systems in which the customers have deadlines, drawn randomly from a general distribution. The performance measure is the fraction of customers arriving to the system that meet their deadline. The earliest deadline first queue discipline is used to maximize performance. The performance of two systems, one in which all customers are served to completion and one in which late customers are dropped, are fully characterized and compared under heavy traffic conditions. The system workloads are approximated by Brownian motion processes with one and two boundaries respectively, and the formulas for the fraction of late customers are derived.

Analytical Challenges in Defense Test and Evaluation

✱Ernest A. Seglie, DOT&E OSD, Director, Operational Test and Evaluation, 1700 Defense Pentagon, Rm 3D947, Washington, DC, DC 20301-1700, *Ernest.Seglie@osd.mil*

Key Words: test, evaluation, policy, DOD, DOE

In the Department of Defense two different offices oversee test and evaluation of weapons systems - a Developmental Test and Evaluation office (DT&E) and an Operational Test and Evaluation (OT&E) office. Historically, developmental test occurred first and operational testing was viewed as a "final exam." Recent changes to T&E policy in December 2007 will require greater cooperation between the offices, greater coordination between the testing each oversee, and possibly the application of new methods and approaches. Among the areas where change is expected is the current tendency of OT to concentrate on 1. fixed configuration testing, 2. hypothesis testing assumptions to determine the size of tests, 3. exclusion of sources of data and information (e.g., DT and modeling results) other than the final OT. These changes must consider political, budgetary, schedule and other constraints.

183 Collecting Spatial Data

Section on Statistics and the Environment, Section on Physical and Engineering Sciences, Section on Survey Research Methods

Monday, August 4, 2:00 p.m.–3:50 p.m.

Spatial Multipurpose Designs

✱Werner G. Mueller, Johannes Kepler University, Altenbergerstr. 69, Linz, 4040 Austria, *werner.mueller@jku.at*

Key Words: compound design, random field, universal kriging

A single purpose design may be quite inefficient for handling a real-life problem. Therefore, we often need to incorporate more than one design criterion and a common approach is simply to construct a weighted average, which may depend upon different information matrices. Designs based upon this method have been termed compound designs. The need to satisfy more than one design criterion is particularly relevant in the context of random fields. It is evident that for precise universal kriging it is important not only to efficiently estimate the spatial trend parameters, but also the parameters of the variogram or covariance function. Both tasks could for instance be comprised by applying corresponding design criteria and constructing a compound design from there. Modern techniques for such first and second order characteristics will be suggested and reviewed in the presentation.

On the Effect of Collocation on the Quality of Multivariate Spatial Prediction

✱Dale Zimmerman, The University of Iowa, 2141 Brown Deer Road, Coralville, IA 52241, *dale-zimmerman@uiowa.edu*

Key Words: Co-kriging

Inferences for spatial data can be affected substantially by the spatial configuration of the network of sites where measurements are taken. Most network design literature considers univariate spatial inference problems, such as kriging and variogram estimation. This presentation considers one aspect of network design for multivariate spatial inference. In particular, I examine the effect that the proportion of collocated design sites has on multivariate spatial prediction, and I show how this effect depends on the relative strengths of spatial correlation and cross-correlation.

Sampling Strategies for Estimating the Spatial Mean of Temporal Trends

✱Dick Brus, Wageningen University and Research Centre, Alterra, PO Box 47, Wageningen, International 6700 AA The Netherlands, *dick.brus@wur.nl*

Key Words: design-based method, model-based method, monitoring, sampling

This paper describes and evaluates two basic approaches for estimating the Spatial Mean of the Temporal Trend: 1. a pure model-based approach with purposive sampling in space and in time, and model-based inference; 2. a mixed approach with probability sampling in space, and purposive sampling in time. In the mixed approach the statistical inference depends on the pattern of the observations in the space-time universe. When sample locations are revisited in all sampling rounds, a first option is to estimate first temporal trends at the sample locations by model-based inference, followed by design-based estimation of the spatial mean of these location-specific temporal trends. A second option is to estimate first the spatial means at the sampling times by design-based inference, followed by model-based inference of the temporal trend of these spatial means.

Spatial Network Design To Detect Regional Trends in PM2.5

✱Zhengyuan Zhu, The University of North Carolina at Chapel Hill, Department of Statistics and Operations Research, CB#3260, UNC-CH, Chapel Hill, NC 27599, *zhuz@email.unc.edu*

Key Words: kriging, optimal design, Bayesian prediction

One goal of the Clean Air Act Amendments of 1990 (CAAA) is to reduce ambient concentrations of atmospherically-transported pollutants. Monitoring data from EPA networks can be used to estimate regional trends of these pollutants to evaluate the effectiveness of the CAAA. This paper presents spatial network design methodology to optimize the network's ability to detect and quantify future regional trends in air pollution by adding or relocating monitoring sites. The 2000–2005 PM2.5 data in the Midwest region of US is analyzed to illustrate the design methodology.

Spatial Sampling Design Using Support Vector Machines

✱Mikhail Kanevski, Institute of Geomatics and Analysis of Risk, IGAR, UNI Lausanne, Amphipole, Lausanne, 1015 Switzerland, *Mikhail.Kanevski@unil.ch*; Alexei Pozdnoukhov, Institute of Geomatics and Analysis of Risk

Key Words: spatial sampling design, support vector machines, active learning

Support Vector Machines (SVM), being robust non-parametric classifiers, have gained considerable success in many data modeling fields. Based on Statistical Learning Theory, they provide a well-founded framework for extracting non-linear dependencies from empirical data in high dimensional spaces. This study develops a novel method for spatial sampling design using SVM for spatial classifications which cover a wide range of tasks in mapping of categorical variables. It was observed, that optimal solutions provided by SVM are often sparse: a larger part of data does not contribute to the decision, while only the support vectors, contribute to describe the decision boundaries between the classes. SVM algorithm gives priorities to new sampling points which the most contributes to the decision process. Spatial sampling design using SVM can be considered as an active learning process.

184 Development and Validation of Biomarkers ●▲

WNAR, Section on Statistics in Epidemiology, Biometrics Section

Monday, August 4, 2:00 p.m.–3:50 p.m.

Imaging Endpoints for Clinical Trials: The Sharp Score for Radiographic Data in Rheumatoid Arthritis

✱Grace S. Park, Amgen, Inc, One Amgen Center Dr., Mailstop 24-2-A, Thousand Oaks, CA 91320, *gpark@amgen.com*

Key Words: Biomarker, Rheumatoid Arthritis, Radiograph, FDA

Imaging is an important tool used to assess and accelerate the process of therapeutic options in clinical medicine. Sharp's criteria for rheumatoid arthritis form just one of three licensed imaging surrogate biomarkers approved by the US Food and Drug Administration (FDA). Along with imaging used in oncology research, the Sharp score serves as a biomarker that reveals the extent and severity of disease using standard radiographs. However, in response to drug development programs and advancements in imaging techniques, the FDA issued its critical path initiative in 2004. Seeking to advance biomarker technologies, it promoted the integration of biomarker evaluation with ongoing trials to reduce development expenses. This presentation gives an overview of the precedent of Sharp's criteria in relation to the scope of the FDA's modernization initiative.

Evaluation Process of Imaging Biomarker: Quantitative Lung Fibrosis (QLF) Score for Computer Tomography (CT) Data in Interstitial Lung Disease

✱Hyun Jung (Grace) Kim, University of California, Los Angeles, 924 Westwood Blvd ste 650, Los Angeles, CA 90024, *gracekim@mednet.ucla.edu*; Gang Li, University of California, Los Angeles

Key Words: Biomarker, Interstitial Lung Disease, CT, Denoise, Classification, image

Imaging modality can vividly show the treatment efficacy in therapeutic intervention. We need to develop and evaluate the vocabulary, imaging biomarker, for the treatment efficacy. The important factors in biomarker evaluation are calibration, discrimination and accuracy in multicenter clinical trial. To be qualified for imaging surrogates biomarker, many challenges are present in the calibration before applying to clinical trial and research application, because different technical factors create a different level of noise and affect to the measurement of imaging biomarker. We present a novel QLF, "biomarker," of CT imaging: denoise as part of calibration and comparison of radiologist score with QLF as discrimination and accuracy. We, then, show this imaging "biomarker," QLF, as outcome variable for the treatment efficacy in the proven scleroderma-lung-disease clinical trial.

Statistical Issues in Discovery and Validation of Genomic Biomarkers for Medical Prognosis and Drug Development

✱MyungShin Oh, Baxter Healthcare Corp., One Baxter way, Westlake Village, CA 91362, *adumas@ucla.edu*

Key Words: Genomic biomarker, Validation, Normalization, Gene Expression Index, Microarray

Over the last decade, innovations in genomic technology have led to an explosion of biomarkers discovery for medical diagnosis and drug development. Biomarker discovery entails the identification of putative markers for the response of interest, typically from exploratory focused multi-array and other technology platforms. This is followed by the process of qualification of these biomarkers via additional studies/cohorts to demonstrate their predictive ability for the intended purpose. To successfully meet these goals, statistical techniques necessary for analyzing experimental results are still being developed, while the specialty of genomic experiments continues to introduce many challenging statistical topics. This presentation will detail some of these statistical challenges, including data pre-processing, normalization, gene expression index, calibration, and clinical validation.

Enrichment Design and Statistical Inference for Cancer Biomarker Validation Study

✱Xiaofei Wang, Duke University Medical Center, Department of Biostatistics & Bioinformatics, Durham, NC , *xiaofei.wang@duke.edu*; Haibo Zhou, The University of North Carolina at Chapel Hill

Key Words: Cancer Research, Biomarker Validation, Enrichment Design, ROC Curve, Semiparametric Inference

In cancer research, there is a growing need in finding biomarkers (e.g., molecular, genetic, imaging) that predict patients' prognosis or response to a specific treatment. Validation of the clinical utility of such biomarkers often require large prospective study. However, due to budget constraint and high cost of bioassays, investigators often have to select a subset from all registered subjects for biomarker assessment. Coupled with the losing of patients with biomarker measures, the correlation between the biomarker and the outcome is commonly moderate and the prevalence rates of positive outcome and positive biomarker are low in the study population. In these situations, investigators face a critical issue of how to select a fixed number of subjects that maximizes statistical power in detecting the correlation of the biomarker and the outcome.

185 Responding to New Challenges: Statistics in Clinical and Translational Science

Biometrics Section, Biopharmaceutical Section, Section on Teaching Statistics in the Health Sciences, WNAR, Section on Health Policy Statistics
Monday, August 4, 2:00 p.m.–3:50 p.m.

Responding to New Challenges: Statistics in Clinical and Translational Science

✱Knut M. Wittkowski, The Rockefeller University, Center for Clinical and Translational Science, 1230 York Ave Box 322, New York, NY 10065, *kmw@rockefeller.edu*; ✱J. Philip Miller, Washington University School of Medicine, *jphilipmiller@wustl.edu*; ✱Frank E. Harrell, Jr., Vanderbilt University, Department of Biostatistics, 1161 21st Avenue South S-2323 MCN, *f.harrell@vanderbilt.edu*; ✱Sally Thurston, University of Rochester, *thurston@bst.rochester.edu*

Key Words: CTSA, biostatistics, epidemiology, research design, training, education

Clinical and Translational Science Awards (CTSAs) support a new initiative funded by the National Institutes of Health. Within the consortium a workgroup of members from each CTSA will focus on methods to integrate biostatistics, epidemiology, and research design (BERD) into clinical and translational science research programs within and across CTSA institutions; encourage the identification and implementation of best practices relating to BERD among institutions; and identify and develop plans to overcome existing barriers to major gains in efficiency in the national clinical research and training efforts. Panel members will briefly present approaches being implemented for a variety of CTSA responsibilities, partnerships, and challenges. Discussion with the audience will follow to explore the extent to which it makes sense to harmonize, communicate, and share across institutions.

186 Small-Area Estimates ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Monday, August 4, 2:00 p.m.–3:50 p.m.

The False Discovery Rate in ACS: Helping Users Understand Estimates for Small Domains

✱Robert E. Fay, Westat, 7252 Greentree Rd, Bethesda, MD 20817, *bobfay@hotmail.com*

Key Words: American Community Survey, Small area estimation, Multiple testing

In 2010, the U.S. Census Bureau will publish the first set of 5-year period estimates from the American Community Survey (ACS), based on data for 2005-2009. Published for small places, tracts, and other small areas, the 5-year ACS estimates will attempt to replace the long-form data from recent decennial censuses. Because the effective sample size for the ACS will be somewhat less than half that of previous censuses, users will face an increased challenge to distinguish true variation from sampling error. The concept of the false discovery rate has become increasingly useful in other disciplines confronted by large numbers of estimates, such as micro-array analysis in genetics and fMRI studies of the brain. The paper will review this concept and suggest its possible future application, based on a preliminary analysis of published data from the ACS Multiyear Estimates Study.

Spatial Modeling and Prediction of County-Level Employment Growth Rate Data

✱Nadarajasundaram Ganesh, National Opinion Research Center, 4350 East-West Highway, Suite 800, Bethesda, MD 20814, *ganesh-nada@norc.org*

Key Words: Fay-Herriot model, Infill asymptotics, Increasing domain asymptotics, Prediction, Spatial Statistics, Small Area Estimation

For correlated sample survey estimates, a linear model is proposed using a covariance matrix in which small areas are grouped into "clusters" by a similarity measure based on spatial locations and covariates. In the context of correlated data, a novel asymptotic framework (a hybrid of infill asymptotics and increasing domain asymptotics) is introduced. The proposed asymptotic framework assumes that the number of "clusters" and the number of small areas in each "cluster" grows with sample size. Under the previously mentioned asymptotic framework, it is theoretically shown that the proposed least squares estimators are consistent and asymptotically normal. Moreover, the theoretical properties of the parameter estimators and robustness to the clustering assumption are verified through simulation. The proposed model is implemented for a county-level civilian employment growth rate data set.

Small-Area Estimation for Alcohol Drinking Among Teenagers

✱Hsing-Yi Chang, National Health Research Institutes, #35, Keyan Road, Zhunan twon, 350 Taiwan, *hsingyi@nhri.org.tw*; Yu-Wen Wen, National Health Research Institutes

Key Words: generalized linear mixed effects model, NHIS, alcohol drinking, teenagers

In this study, we used a two-stage method to estimate the proportion of alcohol drinking among teenagers. Data for this study was the 2001 National Health Interview Survey in Taiwan. We first used generalized linear mixed effects model including the random intercept to estimate the probability of alcohol drinking among teenagers for townships or cities in Taiwan. The estimated random intercept was related to some spatial variables. Then, we used the spatial model to predict the value of the random intercept in the area, where no sample was taken, and their probabilities of alcohol drinking. The results of model show that the highest regional probability of alcohol drinking was about 0.3 in mountainous areas. The model provides reasonable estimates.

Benchmarked Hierarchical Bayesian Posterior Predictive Model Selection

✱Lu Lu, Center for Survey Statistics and Methodology, Iowa State University, 2625 N. Loop Dr., Bldg. 2, Suite 2140, Ames, IA 50011-1272, *icyemma@iastate.edu*; Michael D. Larsen, Iowa State University

Key Words: small area estimation, hierarchical Bayesian, model selection, posterior predictive check, benchmarking, graphical examination

In a stratified multi-stage sample survey conducted by Iowa's State Board of Education, given the budget restrictions, a small sample of PSUs was selected from a two-way stratification of school districts. A hierarchical Bayesian (HB) estimator is proposed to produce more reliable estimates of small area quantities than the design-based direct estimator. A method which benchmarks the HB estimates to the higher levels direct estimates and measures the relative inflation of posterior MSE due to benchmarking in posterior predictions is developed to evaluate the performance of hierarchical models. A graphical examination of posterior predictive discrepancy measures is used. The benchmarked HB posterior predictive model comparison method is shown to be able to select proper models effectively in a simulation study and is then applied to the actual data from ISBE 2004–2005 transcript survey.

Further Developments in a Hierarchical Bayes Approach to Small-Area Estimation of Health Insurance Coverage: State-Level Estimates for Demographic Groups

✱Steven Riesz, U.S. Census Bureau, *steven.riesz@census.gov*; Mark Bauder, U.S. Census Bureau; Donald M. Luery, U.S. Census Bureau

Key Words: administrative data, multivariate modeling, low-income

Fisher et al. (2006) developed a hierarchical Bayes model to estimate the number of people without health insurance within demographic groups for states. The Centers for Disease Control and Prevention are interested in estimates of women without health insurance by demographic groups in families that earn less than 200% of the poverty line. Our approach jointly models survey estimates from the Annual Social and Economic Supplement to the Current Population Survey and estimates from tax, census, food stamp, and Medicaid data using a multivariate, hierarchical approach. We have made important enhancements to the models by including additional administrative data, further elaborating the expectation and variance models for both the survey estimates and the administrative data, and developing an approach to reflect the variation due to raking to survey estimates in the variance estimates.

The Connection Between Bayesian- and Resampling-Based Inference in Small-Area Models

✱Snigdhansu Chatterjee, The University of Minnesota, 313 Ford Hall, 224 Church Street S.E., Minneapolis, MN 55455, *chatterjee@stat.umn.edu*

Key Words: Small area, parametric bootstrap, Bayesian

A general small area model is a hierarchal two stage model, of which special cases are mixed linear models, generalized linear mixed models and hierarchal generalized linear models. Such models naturally lend themselves to both Bayesian and non-Bayesian analysis. In recent times, parametric bootstrap for small area models has become an active field of study. Although resampling techniques like the parametric bootstrap are inherently non-Bayesian, there are some deep connections between resampling and Bayesian analysis. This talk will focus on some of these connections in the context of small area problems.

187 Sampling Frames ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Monday, August 4, 2:00 p.m.–3:50 p.m.

GeoFrame: A Technological Advancement in Field Enumeration

✱Joseph P. McMichael, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709, *mcmichael@rti.org*; Leslie Athey, RTI International; Brian Evans, RTI International; Victoria Albright, RTI International

Key Words: field enumeration, global positioning system (GPS), sampling frame, counting and listing, Google Maps API

Face-to-face interviewing using area probability samples remains the gold standard method for conducting representative surveys, but the increasing cost of these studies has led researchers to look for economies in all phases of survey design and execution. GeoFrameTM is an innovative use of digital photography and geospatial technology that reduces the cost of field enumeration to make construction of area probability samples more affordable. GeoFrame'sTM utility can be extended to other situations where sampling frames may be difficult or too costly to create, such as when creating probability samples to evaluate disaster recovery efforts or conduct surveys in developing nations. This paper describes the initial application of this new technology to construct a sampling frame for a survey of tobacco use in the colonias in El Paso County, Texas.

Using Digital Imagery To Update the Measures of Size of Area Segments

✱Lee Harding, Westat, 1650 Research Blvd., Rockville, MD 20850, *LeeHarding@Westat.com*; Sylvia Dohrmann, Westat; Lin Li, Westat; Lester R. Curtin, Centers for Disease Control and Prevention

Key Words: Area sample, MOS, Probability proportional to size sampling

Two-phase sampling can be used to update the measures of size (MOS) used for selecting area segments when census MOS have become outdated. In the first phase a large sample of segments is selected with probabilities proportional to census MOS and updated MOS are produced. A second phase subsample of segments is then selected making use of the updated MOS. One approach for updating the MOS is to have field staff tour the segments to count the numbers of housing units. The advent of mapping products like Google Earth makes possible an alternative, less expensive, approach that involves counting housing units in the office, based on high-resolution digital imagery. We evaluate the effectiveness of this new approach by comparing field counts, counts from digital imagery, and counts of listed housing units for several high growth PSUs in the National Health and Nutrition Examination Survey.

Success Using an Age-Targeted List Sample for the National Immunization Survey - Adult

✱Steven Pedlow, National Opinion Research Center, 55 E. Monroe, Suite 2000, Chicago, IL 60603, *pedlow-steven@norc.uchicago.edu*; Hee-Choon Shin, National Opinion Research Center; Ashley Amaya, National Opinion Research Center; Gary L. Euler, National Center for Immunization and Respiratory Diseases; Meena Khare, National Center for Health Statistics

Key Words: Vaccination Coverage, Representative Sample, NIS-Adult

The NIS-Adult, a nationwide telephone survey conducted by NORC for the Centers for Disease Control and Prevention, was designed to monitor vaccination rates for persons aged 18 years and older. The initial sample for the NIS-Adult was selected from adult respondents to the National Health Interview Survey (NHIS) who completed the household interview and provided contact information. The NHIS sample had an insufficient number of older adults. To supplement adults aged 50-64 years and 65+ years, an age-targeted list sample based on directory-listed telephone numbers was used. The age-targeted list contributed a larger fraction of the final sample than expected-59%. In this paper, we compare the demographic profiles of the NIS-Adult with those from the Current Population Survey and the NHIS and show that the age-targeted list sample performed above our expectations.

Challenges and Methods in Creating Secondary Sampling Units for Area Probability Samples

✱Bryce Johnson, Westat, 1650 Research Blvd., Rockville, MD 20850, *BryceJohnson@Westat.com*; Jill Montaquila, Westat

Key Words: Secondary sampling units, automation, TIGER, invisible boundaries, contiguous, heterogeneity

In area probability samples secondary sampling units (SSUs), or segments, are often constructed using census blocks, block groups, or tracts. There are numerous challenges in forming segments that meet statistical, operational, and study goals. Creating segments which are composed of contiguous units is often a sought after goal which can be a challenging task whether a manual or automated solution is used. A more complex approach may involve forming segments which are more heterogeneous based on one or several factors. There may also be some advantage in creating more compact segments. An important operational consideration is to create segments with visible boundaries, roads for example instead of property lines. All of these challenges are explored in this paper and a multifaceted solution is presented which relies on automated techniques.

Assessing the Filter Rules for Extracting Addresses from the Master Address File To Construct a Housing Unit Frame for Current Demographic Surveys

✱Joel Martin, U.S. Census Bureau, *joel.m.martin@census.gov*; Clifford Loudermilk, U.S. Census Bureau

Key Words: Sampling Frame, Filter, Surveys, Master Address File

The Master Address File (MAF) is a national inventory of addresses for living quarters in the United State that the Census Bureau continually updates. One of the goals of the 2010 Demographic Household Survey Redesign is to switch to using a MAF-based frame for current demographic household surveys. To create the best possible current household survey frame from the MAF, a set of filter rules must be established to determine which MAF records should be accepted into the frame. The quality of the filter can have a major impact on the coverage of the resulting frame. We assess the effectiveness of various filter criteria by comparing the resulting frames with a benchmark address list collected from the field listings of a nationally representative probability sample.

Evaluating the American Community Survey as a Potential Sampling Frame for the National Survey of College Graduates

✱John Finamore, U.S. Census Bureau, *john.m.finamore@census.gov*; David Hall, U.S. Census Bureau

Key Words: ACS, NSCG, Sampling Frame, Effective Sample

The NSCG is a longitudinal survey that collects information on employment, educational, and demographic characteristics of scientists and engineers in the U.S. The current NSCG sample was selected from the Census 2000 Long Form. In 2010, the NSCG plans to refresh its sample to address attrition and coverage concerns. With the ACS replacing the Long Form, the Census Bureau is evaluating using the ACS as a sampling frame for the NSCG. The evaluation will initially examine the change in available and effective sample if the ACS is used as the sampling frame under the current NSCG design. The evaluation will then examine the potential increase in sampling efficiency for the NSCG associated with the additional demographic information available on the ACS. This paper provides an overview of this evaluation and presents recommendations for the 2010 NSCG based on the evaluation results.

Stratification and Allocation for the Estimation of a Complex Statistic with Auxiliary Data

✱Serge Godbout, Statistics Canada, 150 Tunney's Pasture Driveway, RHC Building, 11G, Ottawa, ON K1A 0T6 Canada, *serge.godbout@statcan.ca*

Key Words: Stratification, Sample allocation, Taylor linearization, Complex statistic

In literature, stratum creation, take-all unit identification and sample allocation methods usually try to optimise the variance or the cost for the estimation of a total. For a complex statistic like a ratio or a regression coefficient, a sampling plan based on the total of one of the variables in the statistic isn't necessarily optimal. We propose to use a method built on a Taylor linearization of the complex statistic. The statistic is first linearized and then, the common stratification and allocation methods can be used. We will measure the efficiency of the method through a simulation study.

We will show an example of the application using the Canadian Survey of Employment, Payrolls and Working Hours.

188 Variations on Designs ●

Section on Statistics in Epidemiology, Section on Survey Research Methods, Biometrics Section
Monday, August 4, 2:00 p.m.–3:50 p.m.

Sampling Issues in Biomarker Studies of Nevus Phenotypes

✱Jaya M. Satagopan, Memorial Sloan-Kettering Cancer Center, 307 East 63rd Street, New York, NY 10021, *satagopj@mskcc.org*

The etiologic effects of complex traits are often conferred by multiple risk factors. Cancer epidemiology studies typically establish a large cohort to study risk factors for complex diseases and disease phenotypes. While some putative risk factor data are obtained from the full cohort, new risk factors such as biomarkers need to be constantly examined as studies continue. Evaluating the biomarkers on the full study cohort may not always be feasible due to logistic or economic concerns. Therefore, it may be useful, or sometimes necessary, to identify an informative sub-cohort for evaluating new risk factors in an efficient manner. This talk examines a novel stratification scheme for identifying an informative sub-cohort for biomarker evaluation. Our work is motivated by the ongoing Study of Nevi in Children the investigates risk factors associated with nevogenesis and nevus evolution.

Assessing the Impact of Intervention

✱Shailendra N. Banerjee, Centers for Disease Control and Prevention, 5375 Country Lake Court, Lilburn, GA 30047, *snb1@cdc.gov*

Key Words: Intervention, Segmented Regrssion, Auto-Correlated, ARIMA

Increasingly, educational, administrative and policy interventions are being carried out to improve public health, quality of medication use and/or contain costs. Several statistical methods have been suggested to implement the intervention procedures and to assess its impact. These methods may vary from randomized controlled trials to Auto-Regressive Integrated Moving Average (ARIMA) to segmented regression analysis or parameter-driven models in regression analysis. Clearly, the methodology is partially guided by nature and availability of data. Thus, ARIMA is more suitable for data with serial dependence. But, it also requires a large number of data points for better results. Segmented regression assumes linearity which often may hold only over short intervals. This method may also not be appropriate for highly auto-correlated data. In this paper, these methods will be compared.

Noncompliance-Corrected Effect of Randomized Highly Active Antiretroviral Therapy on Incident Aids or Death Using Inverse Probability-of-Censoring Weights

✱Lauren E. Cain, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St. E7133, Baltimore, MD 21205, *lcain@jhsph.edu*; Stephen R. Cole, Johns Hopkins Bloomberg School of Public Health

Key Words: Compliance, HIV/AIDS, Inverse probability weight, Trial, randomized

In 1996–7, AIDS Clinical Trial Group study 320 randomized 1156 HIV+ US patients to combination antiretroviral therapy (ART) or highly active ART with equal probability. In one year of follow up, 96 patients incurred AIDS or died, 51 dropped out, and 290 dropped out or stopped their assigned therapy. Noncompliance likely results in a null-biased intent-to-treat hazard ratio (HR) of AIDS or death comparing highly active ART to combination ART: which was 0.75 (95% confidence limits [CL]: 0.43, 1.31) for follow up within 15 weeks. Noncompliance correction using inverse probability-of-censoring weights yielded a 63% stronger HR of 0.46 (95% CL: 0.25, 0.85). Weights were estimated conditional on randomization arm and measured baseline and time-varying covariates. These methods may help resolve discrepancies between randomized studies with differing amounts of compliance.

The Exposure-Controlled Retrospective Study: Application to High-Incidence Diseases

✱Peng T. Liu, U.S. Food and Drug Administration, 5100 Paint Branch Pkwy, College Park, MD 20740, *peng.liu@fda.hhs.gov*

Key Words: Alternative Design, High Incidence, Case-Control Study, Sample Size, Relative Efficiency

For measuring the degree of association between a disease and a risk factor, we have developed a new epidemiological design that can serve as an alternative to the conventional design for cohort and case-control studies. When the prevalence rate of disease is greater than the prevalence rate of exposure, this new design is superior in terms of the sample size and of the efficiency of estimate. A “reversed weighting” procedure is adopted for estimating the required sample size. The relative efficiency of the new design in comparison with case-control study results under a wide variety of conditions is also examined.

Semiparametric-Efficient Estimation for Multistage Case-Control Studies

✱Alastair J. Scott, University of Auckland, 38 Princes StreetDepartment of Statistics, 38 Princes Street, Auckland, International 1010 New Zealand, *scott@stat.auckland.ac.nz*; Alan J. Lee, University of Auckland; Chris J. Wild, University of Auckland

Key Words: multi-stage, case-control study, semiparametric efficiency

In a simple stratified case-control study, a prospective cohort is stratified according to some variables known for the whole cohort. Separate samples of cases and controls are then drawn from each stratum and values of covariates are obtained. In a two-stage study, some of the more expensive or difficult covariates are not measured on all the sampled units, but only on a subsample drawn from them. This process can be continued indefinitely. We derive a set of estimating equations for the efficient semiparametric maximum likelihood estimator with an arbitrary number of stages. These take the form of penalized “pseudo-likelihood” equations, with a term corresponding to an ordinary prospective likelihood plus an extra penalty term for each additional stage of sampling. We illustrate with a three-stage example.

A Steady-State Effect of a Cancer Screening Intervention

✱Shih-Yuan Lee, The University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109, *shihylee@umich.edu*; Alexander Tsodikov, The University of Michigan

Key Words: screening, cancer incidence, prostate cancer

Introduction and increased utilization of cancer screening induces remarkable transient effects on cancer incidence and mortality. When the dust settles we are interested in estimating the achieved benefits comparing a steady state after the intervention with the one before the intervention. We propose methodology for addressing this problem. A model of population under screening is built using methods of queuing theory. The disease free stage and the preclinical disease stage represent two series-connected servers. Stationary states of the system before and after the intervention are derived and linked by a regression model with intervention process characteristics as covariates. The methods are applied to Surveillance, Epidemiology and End Results (SEER) data.

Estimation in Hidden Population Using Social Network

✱Yang Zhao, University of Regina, College West 307.14, Regina, SK S4S 0A2 Canada, *zhaoyang@uregina.ca*

Key Words: Hidden population, Markov chain model, Population proportion, Respondent-driven sampling, Social network relationships

Characteristics of hidden populations (e.g., population of injection drug users) cannot be studied using standard sampling and estimation procedures. This article considers methods for estimating the population proportion of hidden population using social network. We describe an one step sampling procedure for collecting data from both the population and its social network, and provide a method to estimate the population proportion efficiently. We further drive a formula to compute an estimate of the variance of the proposed estimator. Simulation study is provided to illustrate the new sampling and estimation methods.

189 Applications in Statistical Consulting ●

Section on Statistical Consulting

Monday, August 4, 2:00 p.m.–3:50 p.m.

Strategies for New Statisticians Venturing into the Real World

✱Laura Vazquez, Shepherd Center, Marcus Community Bridge Program, 2020 Peachtree Road, NW, Atlanta, GA 30309, *laura_vazquez@shepherd.org*

Key Words: statistical consulting, career development, nontraditional, Masters level

Statistical consulting interactions with colleagues can be rewarding; however, they are often times challenging, difficult and ineffective. Difficult interactions may lead one to consider ethics in statistics, the mistrust of statisticians, evaluate the limits of statistics, both personal and professional, and consider best practices. This presentation will provide strategies for facilitating successful interactions in a clinical environment in which clinicians and researchers evaluate evidence independently and as a part of a team. Strategies will include those learned as an epidemiologist trained at the Masters level, as well as those from other resources, including the American Statistical Association. In addition, the multiple roles of the statistician, along with opportunities for team work in a clinical setting will be explored.

A Total Health Care Costs Analysis for Multiple Sclerosis Patients: Lessons Learned, or How Could My Study Have Been More Efficient?

✱Andrea McAllister, Highmark Inc., 120 Fifth Avenue Place, Suite P7205, Pittsburgh, PA 15222, *andrea.mcallister@highmark.com*

Key Words: matched cohort, total healthcare cost, Specialty Pharmacy, risk scores

A total health care cost analysis of Multiple Sclerosis (MS) patients started out as a straightforward study: MS patients are examined pre- and post-implementation of a disease management program by a health insurer's Specialty Pharmacy provider. What should have been a relatively brief study, about two to three months, ended up being a seven month study using very different parameters than at the onset. Initially all patients, the total cohort, were included for analysis and by the end of the study cohorts matched using DxCG© risk scores were compared. For future studies, can all of the intermediate steps be excluded or are they necessary for comprehensive examination of the populations and resulting data? This question will be explored as well as the benefits and drawbacks to using the DxCG© risk scores in this particular study.

Percent Effort vs. Fee-for-Service: Communicating with Investigators About Cost Recovery

✱Richard F. Ittenbach, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue (MLC5041), Center for Epidemiology and Biostatistics, Cincinnati, OH 45229-3039, *richard.ittenbach@cchmc.org*; Francis W. DeAngelis, The Children's Hospital of Philadelphia

Key Words: consulting, fee for service, percent effort, compensation

Although different models exist for providing statistical support to investigators, two approaches tend to predominate when it comes to the fiscal sustainability of a biostatistics unit, percent effort and fee-for-service. Despite the importance of cost recovery mechanisms to a department, many statisticians are uncomfortable with discussions about the financial implications of their work. Those who are comfortable may not fully understand the rationale for specific policies and procedures. The purpose of this presentation is threefold: first, to identify and describe the two predominant models of cost recovery; second, to offer a rationale for combining the two approaches into a single, overarching framework; and, third, to provide a simple strategy that can be used with investigators when discussing cost recovery components of their work.

Innovative Application of Statistics to Studies from Academic Medical Centers

✱Jayawant Mandrekar, Mayo Clinic, 2665 Boulder Ridge Drive NW, Rochester, MN 55901, *mandrekar.jay@mayo.edu*

Key Words: Graeco Latin Square, Factor Analysis, Deming's regression

Academic medical centers offer opportunities to consult on clinical projects that involve novel applications of statistical methods. In this talk illustrative examples from diverse clinical areas will be presented. Specifically, statistical approaches including; 1) Greco- Latin Square design to compare the performance of four liquid crystal displays under typical medical center lighting conditions, 2) factor analysis to evaluate stability of factorial structure of clinical teaching assessments among medical specialties, 3) determination of the predictive utility of continuous versus categorical scaling of duration of post-traumatic amnesia, and 4) Deming's regression to compare performance of various assays in clinical microbiology studies, will be discussed. Findings from these studies directly impact patient care and/or offer cost savings with improved efficiency.

How Are Mathematics GPA and Several Other Factors Related to University Time-to-Graduation?

✱Xiaohui S. Wang, The University of Texas-Pan American, Mathematics Department, The University of Texas-Pan American, Edinburg, TX 78539, *xhwang@utpa.edu*

Key Words: Ordinal responses, multinomial logistic regression, cumulative logit models

University time-to-graduation is treated as a categorical variable with values Year-4, Year-5, Year-6 and Year-more. Our statistical study shows that mathematics GPA is closely related to Cumulative GPA that usually defines the university graduation year. We will use real student records data obtained from the University of Texas-Pan American to illustrate how mathematics GPA and several other factors, such as gender and full/part-time study status, are related to speed of university graduation.

Estimation of the Correlation Coefficient of Left-Censored and Repeated Measures Data

✱Agustin Calatroni, Rho, Inc., 6330 Quadrangle Drive, Suite 500, Chapel Hill, NC 27517, *acalatro@rhoworld.com*; Cindy Visness, Rho, Inc.; Herman Mitchell, Rho, Inc.

Key Words: correlation, left censored, repeated measures

Cytokines are released by cells in the immune system and inform other cells to generate an immune response. Different cytokines are released depending on the specific immune response required. The correlations among cytokine responses are of interest to immunologists. It is common for laboratory assays of cytokine response to be performed in duplicates or triplicates for quality control purposes. These assays normally have a limit below which the cytokine response cannot be measured, resulting in left censoring. To account simultaneously for left censoring and repeated measures, a maximum likelihood method is proposed to estimate the correlation coefficient using SAS © NLMIXED. The approach is illustrated with data from the Urban Environment and Childhood Asthma Study (URECA).

Parameter Selection for SUP Clustering Algorithm

✱Ting-Li Chen, Institute of Statistical Science, Academia Sinica, *tlchen@stat.sinica.edu.tw*; Shang-Ying Shiu, Institute of Statistical Science, Academia Sinica

Key Words: Clustering, Parameter Selection, Image Segmentation

The Self-Updating Process (Chen and Shiu, 2007) is a simple and novel clustering algorithm. In our previous work, we have shown its promising performance compared to other frequently used clustering methods. However, the effect of parameters on the clustering results still required further investigation. In this talk, we will address the issue of parameter selection. There are two parameters r and T to be determined. r is estimated from the empirical distribution. T is chosen to be proportional to r, or is increasing with respect to the time. The clustering performance of this automatic selection and an example of image segmentation will be presented.

190 Bayesian and Likelihood Inference

Section on Statistical Computing, Section on Bayesian Statistical Science

Monday, August 4, 2:00 p.m.–3:50 p.m.

Bayesian Scale Space Analysis of Image Differences

✱Leena A. Pasanen, University of Oulu, Dept. Of Mathematical Sciences, P.O. Box 3000, FIN-90014 UNIVERSITY OF OULU, FINLAND, Oulu, International FIN-90014 Finland, *Leena.Pasanen@oulu.fi*; Lasse Holmström, University of Oulu

Key Words: Image analysis, Bayesian methods, Scale space, SiZer, BSiZer, Applications

We consider the detection of features in noisy images that appear in different spatial scales or resolutions. Our goal is to capture scale dependent differences in a pair of images of the same scene taken at two different instances of time. A new approach is proposed that uses Bayesian statistical modeling and simulation based inference. The method can be viewed as further development of SiZer technology, originally designed for nonparametric curve fitting. A strength of the Bayesian approach is straightforward inference and incorporation of domain specific prior information on the images in question. While we demonstrate the performance of the new method mostly using artificial test images, the approach is believed to have applications e.g. in satellite based remote sensing. Hence, we also include a preliminary analysis of a pair of Landsat satellite images used in forest inventory.

Bayesian Inference for Measures of Model Fit

✱Joseph Retzer, Maritz Research, 2019 E. River Rd., Grafton, WI 53211, *retzerjj@gmail.com*

Key Words: bayesian, stochastic dominance, information theory, regression, contingency table, entropy

Measures of fit such as assortments of R square's and deviance are functions of model parameters. Uncertainty about the parameters induces uncertainty about the fit measures. The common Bayesian practice is to integrate out the parameters. Stronger inference can be obtained by deriving the distribution of the fit measures from it the constituent parameters. Then the model fit can be estimated by the Bayes estimate of fit under a loss function and/or a posterior interval. Such an inference requires stochastic dominance of the measure of fit for the model over the measures of fit for all of its submodels. We provide algorithms for computing, posterior distributions of fit measures for a various statistical problems including linear regression, contingency table analysis, logit analysis, canonical correlation, and linear structural equations.

Bayesian Nonparametric Model for fMRI of the Visual Cortex

✱Raymond G. Hoffmann, Medical College of Wisconsin, Division of Biostatistics, 8701 Watertown Plank Rd., Milwaukee, WI 53226, *hoffmann@mcw.edu*; Nicholas M. Pajewski, Medical College of Wisconsin; Edward A. DeYoe, Medical College of Wisconsin; Daniel B. Rowe, Medical College of Wisconsin

Key Words: non-parametric, point process, Dirichlet prior, fMRI, imaging, spatial density

The fMRI visual field map (VFM) is obtained by using a rotating-expanding visual target to identify the area of the retina that corresponds to activated visual cortex. Since fMRI data is (1) relative rather than absolute and (2) has a degree of noise that may mask the activation, identifying differences in VFMs requires a model that will differentiate changes in the underlying structure from differences due to imaging variability. The VFM produces a non-homogenous, non-isotropic set of points on a disk that includes irregular features like the blind spot. A non-parametric mixture model, using a Dirichlet prior on a space of 2D density functions, will be used to model the VFM under experimental conditions where part of the visual field is masked by a circular wedge. The posterior probability of the difference in the models, will be used to quantify the probable location of the wedge.

Bayesian and Non-Bayesian Approaches to Sample Size Determination

✱Junheng Ma, Case Western Reserve University, 10900 Euclid Ave., Dept. of Statistics, Cleveland, OH 44106, *jxm216@case.edu*; Jiayang Sun, Case Western Reserve University; Joe Sedransk, Case Western Reserve University

Key Words: predictive posterior p-value, power calculation, sample size, images, Multinomial distribution

In developing new imaging diagnostic tools, it is often of interest to compare the quality of the resulting images from the competing tools. Each image is assessed by one or several evaluators who score the images from 1 to 5. An important scientific question is how many images are needed to detect differences in the quality of these images when using these diagnostic tools. Specifically, one is interested in finding the sample sizes needed to detect a difference between the parameters from two multinomial populations. In this paper we develop three Bayesian approaches based on the ideas of posterior predictive p values and two non-Bayesian methods using a chi-square test and a simultaneous confidence interval. Evaluations and comparisons are made using asymptotics, simulated data and a real data application.

Large Gaussian Covariance Matrix Estimation with Markov Structure

✱Xinwei Deng, Georgia Institute of Technology, 765 Ferst Drive NW, School of Industrial and Systems Engineering, Atlanta, GA 30332, *xdeng@isye.gatech.edu*; Ming Yuan, Georgia Institute of Technology

Key Words: Covariance matrix, Markov structure, Sparsity, Shrinkage estimators, Semi-definite program

Covariance matrix estimation for a large number of Gaussian random variables is a challenging yet increasingly common problem. A fact neglected in practice is that the random variables are frequently observed with certain temporal or spatial structures. Such a problem arises naturally in many situations with time series and images as the most popular and important examples. In this paper, we propose shrinkage estimators of the covariance matrix specifically to address this issue. The proposed methods exploit sparsity in the inverse covariance matrix in a systematic fashion so that the estimate conforms with models of Markov structure and amenable for subsequent stochastic modeling. We show that the estimation procedure can be formulated as a semi-definite program and efficiently computed. We illustrate the merits of these methods through simulation and a real data example.

Simulation-Based Visualization of Inference Functions

✱Daeyoung Kim, The Pennsylvania State University, Department of Statistics, 326 Thomas Building, University Park, PA 16802, *dzk123@psu.edu*; Bruce G. Lindsay, The Pennsylvania State University

Key Words: Inference function, Likelihood ratio statistic, Score statistic, Wald statistic, Visualization, Confidence sets

This paper presents a new simulation based methodology to visualize aspects of inference functions of statistical interest. Our simulation based methodology provides values for the parameters that picture the inference function surface around the estimator optimizing the inference function. Once done, one can picture the profile inference function confidence sets for multiple parameters of interest without further complicated optimization. Although this methodology is related to Fisher's concept of fiducial inference, it is here treated as merely a visualization device. Our method uses the same observed data without repeated sampling, which is also the case in Bayesian inference with Markov chain Monte Carlo algorithm. Moreover, our method does not require the specification of a prior distribution for parameters and a burn-in period for MCMC convergence.

Identifying Statistically Significant Clusters

✱Ranjan Maitra, Iowa State University, Department of Statistics, Ames, IA 50011, *maitra@iastate.edu*; Volodymyr Melnykov, Iowa State University

Key Words: EM algorithm, p-value, second-order delta method

A formal hypothesis-testing approach is presented for detecting significant clusters in data. The approach revolves on calculating the p-value of forming an additional cluster at each stage. We illustrate our methodology using model-based clustering as an example. In this setting, we can use the second-order delta method and obtain exact p-values for the change in optimized likelihood with each additional cluster. Experiments on simulation and classification datasets are promising.

191 Advances in Modeling Physical Data ●

Section on Physical and Engineering Sciences, Section on Nonparametric Statistics

Monday, August 4, 2:00 p.m.–3:50 p.m.

Divergence-Based Kernel for Spectrum Classification and Its Applications

✱Tsukasa Ishigaki, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, 106-8569 Japan, *ishigaki@ism.ac.jp*; Tomoyuki Higuchi, The Institute of Statistical Mathematics

Key Words: Divergence-based kernel, classification, kernel method, kernel classifier, acoustic signal

The proper selection of the kernel function is important in order to obtain the high classification performance for the classifier with kernel method. The presentation describes a high-accuracy classification for the spectra using the kernel classifiers (Support Vector Machine and Relevance Vector Machine) with the divergence-based kernels. The proposed method introduces the divergence, which is a metric between two probability distributions, as a kernel function for similarity calculation of two spectra with the appropriate statistical signal processing. We apply the method to three kinds of acoustic signals observed from real systems. The proposed method demonstrates a higher accuracy than popular kernels, such as the polynomial and Gaussian kernels.

Smoothing Spectral Data

✱Tom Burr, Los Alamos National Laboratory, Mail Stop F600, Los Alamos National Lab, Los Alamos, NM 87545, *tburr@lanl.gov*; Scott Garner, Los Alamos National Laboratory; Nicolas Hengartner, Los Alamos National Laboratory; Steve Myers, Los Alamos National Laboratory

Key Words: smoothing, isotope identification, spectral

A smooth fit to low-dimensional functional data often provides a better estimate of the true average response function than does the raw data. Although there is a large literature on smoothing options, comparison of smoothing options has mostly relied on visual inspection of plots. Spectral analysis in the context of isotope identification for detecting nuclear threats provides recent examples for which a very good estimate of the true response is available. This estimate can be compared to the estimate based on various smooth fits to the data, and therefore the performance of smoothing options can be quantified. We describe two such examples, both of which require locally adaptive smoothing for good performance.

Inference for Multi-Piecewise Regression Using the Bootstrap

✱Trevor Craney, Shell E&P, 8534 Iron Tree Lane, Katy, TX 77494, *Trevor.A.Craney@shell.com*

Key Words: changepoint, pants

Inference in ordinary least squares regression is dependent on various assumptions, including normality. When the least squares approach is constrained, as it is in piecewise regression, the standard assumptions can no longer be assumed valid. The bootstrap approach is applicable to creating confidence, or inference, in this case. Further, if the user allows more than one changepoint to be estimated by the data, a suitable stopping method must be decided, which is not unlike the question of how many variables to add in a multiple regression model using a best subsets or forward selection approach. The bootstrap

is shown to still be applicable in the multi-piecewise regression model with various bootstrap approaches to be considered by the analyst.

Splines with Bounded Extrapolation

✱Spencer Graves, PDF Solutions, 333 W. San Carlos St., suite 700, San Jose, CA 953110, *spencer.graves@pdf.com*

Key Words: splines, modeling, extrapolation

Many modeling methodologies produce unacceptably wild predictions with variations in explanatory variables outside the range of the training set. This research considers various ways of managing this problem using splines that are linear or parabolic in the center of the data but are bounded in various ways outside the range of empirical experience. This mimics some of the desirable properties of a random forest, which can be described as a mixture of splines of order 1 (degree 0).

Fast FSR Methods for Second-Order Linear Regression Models

✱Hugh B. Crews, North Carolina State University, Department of Statistics Campus Box 8203, 2501 Founders Drive, Raleigh, NC 27695, *hugh.crews@gmail.com*; Dennis Boos, North Carolina State University; Leonard Stefanski, North Carolina State University

Key Words: False selection rate, Variable selection, Model selection, Response surface

Many variable selection techniques have been developed that focus on first-order linear regression models. In some applications fitting second-order terms can improve predictive accuracy. We describe forward selection algorithms that use the natural hierarchy existing in second-order linear regression models to limit spurious interactions. We develop stopping rules by extending False Selection Rate methodology to these algorithms. We also propose a general method for controlling multiple-group false selection rates. By estimating a separate entry level for first-order and second-order terms, we obtain equal contributions to the false selection rate from each group. Through Monte Carlo simulation, we compare these methods with alternative estimation methods including the LASSO, CART, and MARS. We apply these methods to optimizing response surface experimental designs.

Interpreting Spline Fits via the Coefficients of Local Taylor Approximations

✱Sundar Dorai-Raj, PDF Solutions, 333 W. San Carlos St., suite 700, San Jose, CA 95110, *sundar.dorai-raj@pdf.com*; Spencer Graves, PDF Solutions

Key Words: Smoothing, Least Squares, Polynomial, Free-knot Splines, Minimal Splines

Splines are piecewise polynomials, and it makes sense to interpret each segment as a local Taylor series approximation about (e.g.) the midpoint of the segment. With free-knot splines, analysts could gain useful information from the estimated spline coefficients. We present R functions that makes it easy to translate the estimated coefficients to more a interpretable "Taylor representation."

192 Data Modeling

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Monday, August 4, 2:00 p.m.–3:50 p.m.

SAIPE County Poverty Models Using Data from the American Community Survey

Wesley Basel, U.S. Census Bureau; ✱David Powers, U.S. Census Bureau, U.S. Census Bureau, Rm. 6H120C, Data Integration Division, Washington, DC 20233-8500, *david.s.powers@census.gov*; Brett O'Hara, U.S. Census Bureau

Key Words: poverty, saipe, acs, fay-herriot, census, small area

The U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program produces model-based estimates of income and poverty using data from Census 2000, the American Community Survey (ACS), administrative records and intercensal population estimates. This work assesses SAIPE county poverty models under log rate and log count data transformations. Equivalent rate and count models are described as the basis for comparison, which differ only in their corresponding dependent variables. Scale invariance and homoskedasticity are assessed, and "goodness of fit" comparisons are made in terms of root mean square error (RMSE). The estimation results support use of the log count model.

Model Utility in a Time Series with Interventions: A Case Study

William L. Seaver, The University of Tennessee; ✱Joy Oliver, University of Tennessee, Knoxville, 408 Stokely Management Center, University of Tennessee, Knoxville, Knoxville, TN 37996, *joliver6@utk.edu*; David J. Woehr, University of Tennessee, Knoxville ; Brian K. Griepentrog, Fors Marsh Group

Key Words: time series, change points, forecast horizons

Time series data with change points at the end of a series can make forecasting particularly complicated because these points can disrupt estimations of fit in the holdout sample. However, poor performance with a holdout sample is not a sign that the model is totally invalid. The current study sought to examine the predictive validity of various time series models, and then extended model fit estimations by forecasting with multiple horizons in order to address the complexity of the data. Implications of this study point to the importance of testing multiple origins and horizons for time series in which the effects of change points are difficult to unmask.

Multivariate Analysis of Waiting and Treatment Time in Emergency Departments

✱Abera Wouhib, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *awouhib@cdc.gov*; Myron Katzoff, Centers for Disease Control and Prevention; Meena Khare, National Center for Health Statistics

Key Words: Multivariate Multiple Regression

In Emergency Departments (ED), patients spend more time waiting before being seen by a provider and thereafter until they are discharged. Although volume of patients and size of EDs have direct effect, additional factors such as demographic characteristics, regions and Health insurance status are believed to have influence on time spend in EDs. Multivariate statistical analysis is applied to the 2000–2005 National Hospital Ambulatory Medical Care Survey (NHAMCS) to establish and evaluate this multidimensional relationship. Some continuous variables, such as waiting and treatment time until discharge, and discrete variables such as: causes of visit, demographic characteristics, regions, and health insurance are analyzed to generate the results. Also, the study tests the validity of the models to explain the underlining relationship.

Tracking Consumer Energy Price Change: An Overview of Federal Data Sources and Methodologies

✱Janice Lent, Energy Information Administration, Energy Information Administration, EI-70, 1000 Independence Ave., SW, Washington, DC 20585, *janice.lent@eia.doe.gov*; Joseph Ayoub, Energy Information Administration

Key Words: price index

Among categories of consumer goods, few have greater potential for rapid price change than consumer energy products. Rising energy prices in recent years have shocked markets and induced many consumers to explore alternative energy sources (e.g., wood stoves, alternative fuel-burning vehicles). Several Federal agencies, as well as private companies, collect data on energy production, consumption, and consumer prices. In addition to the Energy Information Administration, the Bureau of Labor Statistics, the Environmental Protection Agency, and the National Agricultural Statistics Service all collect energy-related data. In this paper, we examine the available sources of consumer energy price and consumption data, compare data from different sources, and discuss the data collection and estimation methodologies used to track inflation in consumer energy prices.

Estimating Valid Signatures on a Petition: Power To Choose Between W and A Goodman Type

✱Mark E. Eakin, The University of Texas at Arlington, Dep't. of Informations Systems and Operations Mana, Box 19437, Arlington, TX 76019, *eakin@uta.edu*; Mary M. Whiteside, The University of Texas at Arlington

Key Words: Nonlinear, Sampling, Replication

To estimate the number of unique valid signatures on a petition, the authors have recently shown that a new non-linear estimator, W, is preferably to the more commonly used Goodman-type estimator provided that a signature's being replicated on the petition is independent of the signature's being invalid for other reasons, such as not a registered voter or wrong address. Through an extensive enumeration, this paper investigates the power of the Chi squared test of independence to determine empirically the preferred estimator for a petition with known levels of petition size, sample size, percent replicated signatures, percent invalid signatures, and degree of association between replicated and otherwise invalid signatures.

A Smoothing Approach to Data Masking

✱Yijie Zhou, Merck & Co., Inc, *yijie_zhou@merck.com*; Francesca Dominici, Johns Hopkins University; Thomas A. Louis, Johns Hopkins University

Key Words: Data Masking, Confidentiality, Spatial Smoothing

Individual-level data are often not publicly available due to confidentiality. Instead, masked data are released for public use. However, analyses performed using masked data may produce biased parameter estimates. We propose a data masking method using spatial smoothing. The method allows for varying both the form and the degree of masking by utilizing a smoothing weight function and a smoothness parameter. We investigate for GLM the bias of parameter estimates resulting from analyses using the masked data, and we show that data masking using a smoothing weight function that accounts for prior knowledge on the spatial pattern of exposure may lead to less biased estimates. We apply the method to the study of racial disparities in mortality, and we find that the bias of the association estimate when using the masked data is highly sensitive to both the form and the degree of masking.

193 Bayesian Methods for Missing Data ●

Section on Bayesian Statistical Science, Section on Survey Research Methods, Section on Health Policy Statistics

Monday, August 4, 2:00 p.m.–3:50 p.m.

Bayesian Analysis for the Difference of Two Poisson Rates with Data Subject to Under-Reporting

✱Brandi Greer, Baylor University, Waco, TX 76798, *brandi_greer@baylor.edu*; Dean Young, Baylor University; James Stamey, Baylor University

Key Words: Under-reporting, Double sampling

We derive a closed-form posterior distribution for the difference of two Poisson rate parameters using both data subject to under-reporting and error free data. We use a Monte Carlo simulation study to examine the characteristics of posterior distributions for the rate difference. We also derive a closed-form posterior distribution for the rate difference when using informative priors in the absence of the error free data. We perform a Monte Carlo sensitivity analysis to examine the effects of the priors on the posterior distribution of the rate difference. We then apply our methods to an example comparing automobile accident rates for males and females. We demonstrate a considerable difference in credible sets for the rate difference found using our two Bayesian methods that account for under-reporting versus a Bayesian interval that does not account for under-reporting.

A Bayesian Cox-Regression Model with Informative Censoring

✱Niko Kaciroti, The University of Michigan, 300 N. Ingalls bldg., Ann Arbor, MI 48109, *nicola@umich.edu*

Key Words: Time-to-event analysis, Missing data, Gibbs sampling, Hypertension, Randomized clinical trials

In randomized clinical trials the primary outcome of interest often is time to the occurrence of an event. Cox-Regression models are commonly used to analyze such type of data. The standard Cox model assumes that the censored data are non-informative. These assumptions are usually not testable from the observed data and may well introduce biased. This paper presents a method for analyzing time to event data using Cox-Regression when censoring is informative. We propose a Bayesian model to analyze this type of data by introducing informative prior distributions to identify the model. Sensitivity analysis is then used over a range of these prior distributions. The method is applied to analyze the data from the Trial of Preventive Hypertension Study.

Effects of Missing and Censored Data for Nonlinear Models Involving ODEs

Sujit Ghosh, North Carolina State University; ✱Haojun Ouyang, North Carolina State University, Bioinformatics Program, Raleigh, NC 27695-7566, *houyang@ncsu.edu*

Key Words: Bayesian inference, Censored data, Data augmentation, Missing data, Monte Carlo simulation

The Bayesian Euler's Approximation Method (BEAM) has recently been proposed to estimate the parameters in a non-linear model involving with ODEs, especially when analytical closed form solutions are not available. In this article, the BEAM is extended to handle datasets with missing or censored observations. The proposed method is based on data augmentation algorithm. A simulation study based on growth colonies of paramecium aurelium is presented to compare the performances of the proposed method for various percentages of missing and censored data cases and results are compared to complete data case. Finally the method is illustrated with a real data of AIDS Clinical Trials Group Protocol 315.

Multivariate Mixed-Effects Models for Imputing Panel Missingness in Longitudinal Administrative Data Sets

✱Recai M. Yucel, SUNY-Albany, Department of Epidemiology and Biostatistics, One University Place, School of Public Health, Rensselaer, NY 12144-3456, *ryucel@albany.edu*

Key Words: multiple imputation, Bayesian inference, missing data, multiple membership, mixed-effects, administrative data

Existence of missing values is an inevitable problem facing researchers in many fields. In data systems with complexities due to clustering (both nested and non-nested) and multiple-membership, missing values present analytic challenge in the statistical analyses. I discuss inference multiple imputation (MI) in this context. I specifically focus on panel missingness in administrative longitudinal applications. Adaptations of multivariate generalizations of the mixed-effects models are used to draw inference by MI. Markov Chain Monte Carlo techniques are used to simulate and draw imputations from underlying joint posterior predictive distributions. Brief discussion on handling mixture of variable types and calibration techniques for post-imputation checks will be provided and demonstrated using a longitudinal administrative data.

Bayesian Semiparametric Density Regression with Measurement Error

✱Ju-Hyun Park, The University of North Carolina at Chapel Hill, 5523 NC Highway 55 #434, Durham, NC 27713, *juhyunp@email.unc.edu*; David Dunson, National Institute of Environmental Health Science

Key Words: Conditional density model, Density regression, Dirichlet process, Kernel stick-breaking process, Latent variables, Nonparametric Bayes

In many applications, it is of interest to study the relationship between a predictor X and a response Y without imposing restrictive parametric assumptions on the conditional response distribution of Y given X. The focus of this article is on developing a semiparametric Bayes approach for flexible conditional response distribution modeling, accommodating predictors that are measured with error without imposing parametric assumptions on the distribution of the missing predictor. Our approach relies on a joint modeling strategy, which uses Dirichlet process mixture models for the distribution of X and kernel stick-breaking process mixtures for the conditional distribution of Y given X. Identifiability issues are considered, and an efficient MCMC algorithm is developed for posterior computation. The methods are illustrated using simulation examples and an epidemiologic application.

Bayesian Variable Selection for Large Data Sets with Missing Covariates

✱Jessica Z. Li, University of Florida, Department of Statistics, 102 Griffin-Floyd Hall, P.O. Box 118545, Gainesville, FL 32611, *zhenli79@ufl.edu*; George Casella, University of Florida

Key Words: Bayes factor, missing covariates, updating matrix inversion, stochastic search, Bridge sampling, Bayesian variable selection

Bayesian variable selection has recently become a popular research topic, and has been applied successfully in many situations. However, for the situation of missing covariates, less progress has been made. To find good subsets of variables from large amount of covariates, and at the same time to deal with the problem that a certain percentage of the covariates are missing, we propose a Bayes factor driven search algorithm to find best subsets. With missing data it is difficult to find a closed form for the Bayes factor, so we derive a bridge sampling type approximation. We prove that the stationary distribution is preserved and ergodicity is achieved when the approximate Bayes factors are used in a stochastic search algorithm. To increase the computational speed, we use a matrix inversion identity to accommodate the situation.

194 Estimation, Testing, and Clustering in High Dimensions

IMS, Section on Nonparametric Statistics
Monday, August 4, 2:00 p.m.–3:50 p.m.

Estimating Individual Effect Sizes and Local FDR in Multiple-Hypothesis Testing

✱József Bukszár, Virginia Commonwealth University, 410 North 12th Street, Medical College of Virginia Campus, Richmond, VA 23298, *jbukszar@vcu.edu*; Edwin J.C.G. van den Oord, Virginia Commonwealth University

Key Words: multiple-hypothesis testing, individual effect size, local FDR, genome-wide association

We present a method that estimates the individual effect sizes (IES) in multiple-hypothesis testing problem. A novel and key element is that we use the real likelihood function of the multiple tests rather the one based on the frequently used mixture model. Interestingly enough, this latter one would result in fairly inaccurate estimate even for arbitrarily high sample size. Our IES estimates enable us to estimate the posterior probability that a hypothesis is true null, i.e. the local FDR, accurately. Further applications of IES, such as the estimate of Storey's q-values, will also be discussed. Our methods will be demonstrated on analysis of genome-wide association data.

Exact Calculations of Expected Power for the Benjamini-Hochberg Procedure

✱Deborah Glueck, University of Colorado Denver, Department of Preventive Medicine and Biometrics, B119 Room 2C63, Office Annex 4200 East Ninth Avenue, Denver, CO 80262, *Deborah.Glueck@uchsc.edu*; Jan Mandel, University of Colorado Denver; Anis Karimpour-Fard, University of Colorado Denver; Lawrence Hunter, University of Colorado Denver; Keith E. Muller, University of Florida

Key Words: hypothesis testing, multiple comparisons, false discovery rate, distribution of rejections

Exact analytic expressions are developed for the average power of the Benjamini and Hochberg false discovery control procedure. The result is based on explicit computation of the joint probability distribution of the total number of rejections and the number of false rejections, and expressed in terms of the cumulative distribution functions of the p-values of the hypotheses. We give an example power analysis for a multiple comparisons problem in mammography.

Step-Up Tests for Number of Active Effects in Orthogonal Saturated Designs

✱Samuel S. Wu, University of Florida, 1329 SW 16th St, Gainesville, FL 32610, *samwu@biostat.ufl.edu*; Weizhen Wang, Wright State University

Key Words: Closed test method, Effect sparsity, Experimentwise error rate, Stochastic ordering, Non-central chi-squared distribution

A sequence of null hypotheses regarding the number of negligible effects (zero effects) in orthogonal saturated designs is formulated. Some stochastic ordering results are obtained for functions of normally distributed effect estimates. With these results and the assumption of effect sparsity, we show that two step-up simultaneous testing procedures strongly control experimentwise error rate.

G-PLUS Algorithm for Concave-Penalized Negative Likelihood Selection

✱Wenhua Jiang, Rutgers, The State University of New Jersey, Department of Statistics, 110 Frelinghuysen Road, Piscataway, NJ 08854, *wenhua@stat.rutgers.edu*; Cun-Hui Zhang, Rutgers, The State University of New Jersey

Key Words: LASSO, SCAD, penalized likelihood, solution path, variable selection

Penalized methods are widely used for variable selection in high-dimensional data analysis. Concave-penalized negative likelihood approach like the SCAD has caught much attention since it enjoys promising theoretical properties. In this talk we discuss generalized PLUS (G-PLUS) algorithm to compute the solution path of concave-penalized negative likelihood. The G-PLUS is able to discover the possibly multiple local minimizers of an individual penalty level by continuously tracing the minimizers of other different penalty levels. We use numerous end-to-end short linear segments to approximate the nonlinear paths of generalized linear models. Our simulation results show that both the SCAD and the minimax concave penalty (MCP) methods remarkably outperform LASSO in linear and logistic regression.

Model-Based Clustering with Nonparametric Maximum Likelihood

✱Yeojin Chung, The Pennsylvania State University, 330A Thomas Building, University Park, PA 16802, *ychung@psu.edu*; Bruce G. Lindsay, The Pennsylvania State University; Jia Li, The Pennsylvania State University

Key Words: clustering, nonparametric maximum likelihood, EM algorithm, HMAC, MEM

Clustering analysis is widely used as a major data mining technology. Li, Ray, and Lindsay (2007) developed the Modal EM(MEM) algorithm, searching for a mode of a mixture density starting from any given point. Applying MEM to the kernel density estimate, they derived the Hierarchical Mode Association Clustering (HMAC), grouping data points associated to the same mode. However, the kernel density estimate is naturally biased which could lower the performance of clustering. We investigate the improvement in HMAC by density estimators, based on treating the kernel density estimator as an element of the model consisting of all mixtures of the kernel, continuous or discrete. One can "likelihood tune" the kernel density estimator by using it as the starting value in an EM algorithm and it leads to a fitted density with higher likelihood and smaller bias than the kernel density estimator.

Adaptive Regularization Through Entire Solution Surface

✱Seongho Wu, The University of Minnesota, 313 Ford Hall, 224 Church St. S.E., School of Statistics, Minneapolis, MN 55455, *swu@stat.umn.edu*

Key Words: Homotopy, Least squares, Subdifferential, Support vector machine, Variable selcetion, Variable grouping

Several sparseness penalties have been suggested for delivery of good predictive performance in automatic variable selection within the framework of regularization. All assume that the true model is sparse. We propose a penalty, a convex combination of the L1- and Linf-norms, that adapts to a variety of situations including sparseness and nonsparseness, grouping and nongrouping. The proposed penalty performs grouping and adaptive regularization. In addition, we introduce a novel homotopy algorithm utilizing subgradients for developing regularization solution surfaces involving multiple regularizers. This permits efficient computation and adaptive tuning. Numerical experiments are conducted via simulation. In simulated and real examples, the proposed penalty compares well against popular alternatives.

195 Multivariate Survival Analysis: Theory and Methods

Biometrics Section, Section on Quality and Productivity, Section on Statistics in Epidemiology, WNAR

Monday, August 4, 2:00 p.m.–3:50 p.m.

The Analysis of Bivariate Truncated Data Using the Clayton Copula Model

✱Antai Wang, Georgetown University, *aw94@georgetown.edu*

Key Words: Copula models, Goodness-of-fit test, The Clayton Model, Bivariate truncated data

In individuals infected with human immunodeficiency virus (HIV), distributions of quantitative HIV RNA measurements may be highly left-censored due to values falling below assay detection limits (DL). It is of the interest to find the relationship between plasma and semen viral loads. To address this type of problem, we developed an empirical goodness-of-fit test to check the Clayton model assumption for bivariate truncated data. We also used truncated tau to estimate the dependence parameter in the Clayton model for this type of data. It turns out that the proposed methodology works for both truncated and fixed left censored bivariate data. The proposed test procedure is demonstrated using an HIV data set, and statistical inference is drawn based on corresponding test result.

Semiparametric Estimation with Correlated Recurrent Event Data Under Informative Monitoring

✱Akim M. Adekpedjou, University of Missouri-Rolla, Department of Mathematics and Statistics, 304 Rolla Building, Rolla, MO 65409, *akima@umr.edu*; Jonathan Quiton, Western Kentucky University

Key Words: Correlated Recurrent times, Informative Censoring, Koziol-Green Model, Frailty model

Consider a recurrent event data where frailty models are used to account for correlations among the inter-event times within each unit under study. In this talk we consider the problem of semiparametric estimation of the inter-event time distribution under informative monitoring and the Gamma frailty model. We propose a semiparametric estimator of the baseline survivor function. We show that the estimator under the i.i.d. setting is inconsistent in the presence of frailty. We present results of simulaiton study where we discuss the performence of our proposed estimator to that derived under the i.i.d. setting. Finally, these estimators will be demonstrated by applying to biomedical and engineering data sets.

A Frailty Model in Crossover Studies with Time-to-Event Response Variable

✱Ilya Shvartsman, The Pennsylvania State University Harrisburg, Department of Mathematics and Computer Science, 777 West Harrisburg Pike, Middletown, PA 17057, *ius13@psu.edu*; Vernon Chinchilli, Pennsylvania State College of Medicine

Key Words: crossover design, survival analysis, frailty, maximum likelihood

In this paper, we develop a model to study treatment, period, carryover, and other applicable effects in a crossover design with time-to-event response variable. Since times-to-event on different treatment arms are correlated for each individual, we adopt a proportional hazard frailty model. In the case when frailty is assumed to have a gamma distribution, and hazard rates are constant on each treatment arm (the latter assumption is reasonable for chronic diseased on a relatively short period of time), the likelihood function for determining parameters of interest can be evaluated in a closed form. We apply the developed model to real data collected from asthma treatment studies.

Marginal Hazards Regression for Case-Cohort Studies with Multiple Disease Outcomes

✱Sangwook Kang, The University of Georgia, 129A Paul D. Coverdell Center, Athens, GA 30602, *skang@uga.edu*; Jianwen Cai, The University of North Carolina at Chapel Hill

Key Words: case-cohort study, marginal hazards models, multiple disease outcomes, multivariate failure times, survival analysis, weighted estimating equations

A case-cohort study design is widely used to reduce the cost of large cohort studies, especially when the disease rate is low. A key advantage of the case-cohort study design is its ability to use the same subcohort for several diseases. In order to compare the effect of a risk factor on different diseases, times to different events need to be modeled simultaneously. Valid statistical methods taking the correlations among the outcomes from the same subject into account need to be developed. To this end, we consider marginal proportional hazards regression model and propose an estimating equation approach for parameter estimation. We also consider the generalized case-cohort designs which do not require sampling of all the cases. Asymptotic properties and finite sample properties of the proposed estimators are investigated. The proposed methods are applied to the Busselton Health Study.

Extension of Flowgraph Models with Covariates for Evaluating Effects of Kidney Retransplantation

✱Ya-Hui Hsueh, Tulane University School of Public Health and Tropical Medicine, 1440 Canal St. Suite 2001, New Orleans, LA 70112, *yhsueh1@tulane.edu*; C. Lillian Yau, Tulane University School of Public Health and Tropical Medicine

Key Words: Saddlepoint Approximation, Survival Analysis, time-to-event data, Flowgraph model, semi-Markov models

Flowgraph models offer a useful data analytical tool for multistate models, and have been applied effectively in modeling censored and incomplete survival data. Recently, it has been shown that flowgraph model can also serve as a regression technique to assess the effects of covariates under the Bayesian framework. We further the development of detecting the effects of covariates by maximizing the likelihood function of the flowgraph model. Hypothesis testings for estimated parameters of covariates in a flowgraph model are proposed. Kidney transplant data from United Network for Organ Sharing (UNOS) Organization Registry is used to illustrate the extension of flowgraph models. The proposed model will enhance the ability for evaluating the effects of kidney retransplantation.

Prediction in Multivariate Survival Models

✱M. Brent McHenry, Bristol-Myers Squibb Company, 5 Research Parkway, Wallingford, CT 06492, *brent.mchenry@bms.com*; Stuart R. Lipsitz, Harvard Medical School; Debajyoti Sinha, Florida State University; Sundar Natarajan, VA New York Harbor Harbor Healthcare System

Key Words: Multivariate survival analysis, full-likelihood approach, prediction, oncology

Multivariate survival data arise from time-to-event studies when either of two or more types of events occur for the same subject. In these studies, survival times are correlated within subject, violating the independence of survival times assumption otherwise required. A popular approach for analyzing such data is the marginal approach. We propose a full likelihood approach to estimate the conditional risk, given the history of another event. We evaluate the effect of treatment on the risk of one outcome given the history of other events and covariates. We estimate the joint distribution of the survival times, and the conditional distribution for one survival time given the other survival times. This new approach allows one to predict risk for one particular event, given the past history of other events. The method will be illustrated using an oncology example.

196 Volatility

Business and Economics Statistics Section, IMS
Monday, August 4, 2:00 p.m.–3:50 p.m.

Autoregressive Models with Missing Data and Dependent Innovations

✱Natalia Bahamonde, Pontifica Universidad Católica de Valparaíso, Instituto de Estadística, Blanco Viel 596, Cerro Barón, Valparaíso, 2350026 Chile, *natalia.bahamonde@ucv.cl*

Key Words: time series, missing observations, ARCH models, weak assumptions

In this work, we are interested in parametric estimation in stationary autoregressive model in presence of missing observations. Weak assumptions on the innovations sequence are made since the noise can be a martingale difference sequence with time-varying conditional variance. These assumptions are satisfied for instance in the case of an Autoregressive Conditional Heteroscedasticity model. These models have been introduced by Engle (1982) and are used extensively in economics. We propose a least squares estimator of the parameters of autoregressive models with missing data whose innovations are dependent. The statistical properties of this estimator are explored and the asymptotic normality is derived. The performance of our method is illustrated by numerical examples and applications to real data sets involving series of financial returns of stocks indices from Latin-American markets.

The Contribution of Jumps to Price Volatility in the Energy Futures Market

✱Johan Bjursell, George Mason University, Fairfax, VA , *cbjursel@gmu.edu*; James E. Gentle, George Mason University; George H K. Wang, George Mason University

Key Words: jumps, realized volatility, microstructure noise, energy futures market, theory of storage

We apply a nonparametric method based on realized variations calculated from intraday data to test for jumps in daily futures price series of heating oil and natural gas contracts traded in the New York Mercantile Exchange. The sample period of our intraday data covers January 1996 to December 2005. Alternative methods such as staggered returns and noise filter methods are used to remove the effects of microstructure noise biases on the tests against

detecting jumps. Our empirical work documents the seasonal variation of the frequency of jumps and relative contribution of jumps to total futures price variation. In general, these results are consistent with the implications from the theory of storage.

GARCH Model with Non-I.I.D. Rescaled Errors

✱Kazuhiko Shinki, University of Wisconsin-Madison, 401 N Eau Claire Ave, Apt 310, Madison, WI 53705, *seikibunpu@yahoo.co.jp*

Key Words: GARCH sequence, time series, quasi-maximum likelyhood, consistency, asymptotic normality, random matrices

The GARCH model is widely used as a normalizing procedure for time series, and many studies test the dependence of the resulting rescaled errors. Under such circumstances, it is worthwhile to explore the validity of GARCH estimators under the weakest possible conditions on rescaled errors. This paper proves consistency and asymptotic normality of Gaussian quasi-MLE for GARCH(p,q) models under stationary and ergodic rescaled errors. This result generalizes Lee and Hansen (1994) for GARCH(1, 1) and Berkes, Horvath and Kokoszka (2003a) for GARCH(p,q) with I.I.D. rescaled errors. By extending the methods in the Berkes paper, almost sure and global convergence of a quasi-MLE estimator to the true GARCH parameter is proved. Lastly one applicable probabilistic example is shown, using properties of Markov chains.

Memory Structure in Stochastic Volatility Models

✱Wen Li, Iowa State University, *shirley@iastate.edu*; Cindy Yu, Iowa State University; Alicia Carriquiry, Iowa State University; Wolfgang Kliemann, Iowa State University

Key Words: CIR model, Fractional Brownian motion, Hurst exponent, R/S statistic, Stochastic volatility model

The Black-Scholes model is a continuous-time model that can set the prices of financial derivatives. The model assumes constant volatility, but it has been amply demonstrated that in real markets, the back-out volatility follows a curve. Therefore, more general pricing models with stochastic volatility have been proposed. An example is the CIR model. Little is known about the ability of these models to capture possibly complex memory structures in data. That is, whether the memory structure underlying price and volatility in models captures the memory structure underlying real data is unclear. The focus of this work is on examining whether the CIR model accurately represents the memory structure of real financial price data. We also discuss an innovative form of the CIR model that could capture more complex memory structures than the usual stochastic diffusion processes.

How Potent Are News Reversals? Evidence from the Futures Markets

✱Rohan A. Christie-David, University of Louisville, COB Building, College of Business, Department of Finance, Cardinal Blvd, Louisville, KY 40292, *rohan.christiedavid@louisville.edu*; Arjun Chatrath, University of Portland; Kiseop Lee, University of Louisville

Key Words: News, Reversal, volatility, trading, macroeconomic

We present a theoretical model that predicts a heightening in return volatility following a news reversal. A reversal occurs when a value of an economic indicator that is larger than the forecasted value is followed in the following month by a value smaller than the forecasted value, or vice versa. Our model also suggests that the effects of a news reversal will be more pronounced early in the monthly macroeconomic news cycle. The predictions of the model for trading activity are less clear. We test the main predictions of the model employing intraday data for the nearby Treasury bond futures contract. Consistent with the model, the data show significantly greater responses in volatility when there is a news reversal, than otherwise. Further, the increased sensitivity of volatility is especially perceptible early in the announcement cycle. The findings for trading are less clear.

Empirical Volatility and Seasonality Models for Electricity Futures Prices

✱Hai Dong, University of South Carolina, 29205, *hai_dong@moore.sc.edu*

Key Words: electricity market, forward contract, stochastic volatility, jump diffusion, financial derivatives, term structure

The deregulation in the electricity wholesale industry has lead to a recent boom in the electricity commodity market. However, the lack of an appropriate econometric model to describe the volatility structure and the seasonal characteristic in both spot and forward commodity price processes imposes a major impediment to the newly developed market. In this study, I fit several candidate models to electricity forward prices. My findings suggest that the stochastic volatility model with a jump diffusion component and a seasonality component exhibit certain advantages over the alternatives.

Filtering and Option Pricing with Transformation

✱Hui Gong, Temple University, 203H Speakman Hall, 1810 N. 13th Street, Philadelphia, PA 19122, *hgong01@temple.edu*; Aerambamoorthy Thavaneswaran, University of Manitoba; Jagbir Singh, Temple University

Key Words: filter, option, pricing, transformation

Black-Scholes model is well-known and provides a well defined closed-form formula for option pricing. However, empirical evidence suggests the constant volatility assumption in the Black-Scholes model is not realistic. Using a transformation proposed by Shiryayev (1995), we have studied the linear filtering problem for stochastic volatility models in Gong (2008) et. al. In this paper, we study the option pricing for some continuous time stochastic volatility models using transformation.

197 Statistical Methods in Bioinformatics

Biometrics Section, Biopharmaceutical Section, WNAR

Monday, August 4, 2:00 p.m.–3:50 p.m.

Bayes' Odds Estimation for Protein-Protein Interactions

✱Julia L. Sharp, Clemson University, 237 Barre Hall, Clemson, SC 29634, *jsharp@clemson.edu*; Kevin K. Anderson, Pacific Northwest National Laboratory; Don S. Daly, Pacific Northwest National Laboratory; John J. Borkowski, Montana State University; Gregory B. Hurst, Oak Ridge National Laboratory; William R. Cannon, Pacific Northwest National Laboratory

Key Words: protein-protein interaction, affinity-isolation, Bayes' Odds

Understanding protein complexes is an important aim of proteomics, the large-scale study of protein structures and functions. This understanding may be gained through a "bait-prey" protein affinity-isolation assay that uses a fixed "bait" protein to capture interacting "prey" proteins with LC-MS (liquid chromatography-mass spectrometry) to reveal their identities. LC-MS identifies numerous but not all sample proteins so that a "bait-prey" experiment generates a sparse prey-by-bait score matrix from which a prey-by-bait frequency matrix follows. The posterior probability of a prey-bait interaction is estimated from these frequencies with Bayes' Odds using prior probabilities estimated from prior protein knowledge and descriptive statistics. A study using known complexes of the Rhodopseudomonas palustris bacterium demonstrated the method produced useful, biologically sensible inferences.

Statistical Methods for Genetic Association Study in Admixed Population

✱Qinyi Cheng, New York University, 10016, *qc218@nyu.edu*; Mengling Liu, New York University; Yongzhao Shao, Iowa State University

Key Words: population admixture, haplotype, gene-gene and gene-environment interactions, complex disease, ancestry-informative marker

Population-based genetic association studies have been extensively used to understand the role of genetic variants in complex disease. When the study population is a mixture of subpopulations, ordinary methods neglecting the population structure are prone to spurious associations. The primary objective is to develop statistical approaches for association studies of qualitative and quantitative traits and high-throughput genetic factors when adjusting population admixture using ancestry-informative markers. The population admixture will be incorporated into the analysis of simple genetic structure and haplotype analysis, as well as the study of gene-gene and gene-environment interactions, which are proven to be important in the development of many complex diseases.

Partial Genomic Selection in Animal Populations

✱Natascha Vukasinovic, Newsham Choice Genetics, 800 N. Lindbergh Blvd., Saint Louis, MO 63167, *nvukasin@charter.net*

Key Words: QTL, marker assisted selection, SNP, genome, genomic selection, genetic marker

Marker assisted selection (MAS) has been used to increase accuracy of estimated genetic values of animals by using markers close to the detected quantitative trait loci (QTL). Advances in genome sequencing have produced thousands of SNPs, making genomic selection theoretically possible. In reality, MAS in animals is based on a few known QTL and uses a small number of genetic markers. The costs of routine genotyping limit commercial applications of genomic selection. A partial genomic selection approach can be considered as an alternative. In this approach, genotyping is focused on SNPs in the vicinity of known QTL positions with the largest estimated effects. Effects of other genes are accounted for by single SNPs or the additive polygenic effect. In this study, we examine statistical properties, accuracy, and computational feasibility of the partial genomic selection approach.

A Bayesian Approach To Estimate P(R) Curve from Small-Angle Scattering (SAS) Images

✱Sudeshna Paul, Purdue University, 150 N University Street, West Lafayette, IN 47907, *paul0@stat.purdue.edu*; Alan Friedman, Purdue University; Chris Bailey-Kellogg, Dartmouth College; Bruce Craig, Purdue University

Key Words: indirect transform, scattering intensities, bayesian, correlation

SAS is a useful technology to study the structure of protein in solution. Commonly, SAS experiments determine the inter-atomic distance distribution, or P(r) curve, which represents the maximum derivable structural information. With other approaches, summarized scattering intensities (standard errors) are collected from detector images and indirect transform methods are used to construct P(r) curve. In this talk, we will describe a Bayesian approach to construct P(r) that utilizes the detector images directly. This allows us to extract information from the images, including possible correlation among the pixels and varying image center. It also provides a direct assessment of the uncertainty in P(r). We utilize the indirect-transform method of Moore (1980) but our approach can handle any representation. We will demonstrate this approach using simulated and experimental detector images.

Disentangling the Selection Pressures Acting on Overlapping Reading Frames

✱Karin S. Dorman, Iowa State University, Department of Statistics, Ames, IA 50011, *kdorman@iastate.edu*; Wei-Chen Chen, Iowa State University

Key Words: open reading frame, selection, evolution, continuous time markov chain

We describe a strategy for modeling the evolution of overlapping reading frames, where gene products are coded in different phases by the same genetic sequence. Continuous time Markov models of evolution do not extend easily to overlapping reading frames because codon overlap puts the whole observed sequence in a high dimensional state space without the usual iid replicate observations of sites (nucleotides, amino acids, or codons in existing models). We focus on relatively conserved sequences and identify mutating codon blocks. Assuming these codon blocks are independent, we can test the hypothesis that both reading frames are under the same selection pressure. When applied to sequence data from the Equine Infectious Anemia Virus, we find that the Transmembrane protein is under stronger purifying selection than the Rev protein produced by the same sequence.

Prediction of Protein Inter-Domain Linker Regions by a Nonstationary Hidden Markov Model

✱Kyounghwa Bae, American College of Radiology, Radiation Therapy Oncology Group, 1818 Market st. Suite 1600, Philadelphia, PA 19103, *kbae@phila.acr.org*; Bani K. Mallick, Texas A&M University; Christine G. Elsik, Georgetown University

Key Words: Protein structure prediction, Bayesian model, Inter-domain linker, Non-stationary hidden Markov model, Markov chain Monte Carlo

Proteins have two types of regions: linker regions and non-linker regions which include domains and terminal residues. It is important to try to distinguish these regions using information in the sequence of amino acids that, when folded, form the protein. In this article we develop a nonstationary hidden Markov model to infer region boundaries. The approach is based upon a latent variable defined on the sequence; this variable is a continuous index of the region type and affects the probabilities that specific amino acids will appear. We develop an efficient Bayesian estimate of the model using Markov chain Monte Carlo methods, particularly Gibbs sampling, to simulate the parameters from the posteriors. We applied our method to protein sequences with domains and inter-domain linkers delineated using the Pfam-A database. The prediction results are superior to simpler methods.

198 Issues Related to Clinical Endpoints and Biomarkers ●▲

Biopharmaceutical Section, Section on Statistics in Epidemiology, WNAR, Biometrics Section

Monday, August 4, 2:00 p.m.–3:50 p.m.

Do Opinions from Statisticians on Endpoints Count?

✱Ling Chen, U.S. Food and Drug Administration, 10903 New Hampshire Ave, WO22 RM5241, Silver Spring, MD 20906, *ling.chen@fda.hhs.gov*

Key Words: Multiple co-endpoints, Responder, VVA studies

Most pharmaceutical statisticians make contributions to drug development at the study level by designing and analyzing clinical studies, and believe

that determining primary or secondary endpoints is purely the job of the clinicians. In this talk, I will give two examples to show that statisticians could provide constructive suggestions on endpoints. Statisticians' viewpoints on endpoints may be taken into consideration by the clinicians.

Combining Biomarker Selection and Modeling in Early Clinical Development

✱Alan Y. Chiang, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *chiangay@lilly.com*

Key Words: Biomarker, Variable selection, Phase I/II, Spline

One of the aims of early phase clinical trials is to determine the dosage that will be investigated during a confirmatory clinical trial. Biomarker candidates are generally used as surrogates for clinical endpoints to measure the pharmacological activity of the drug. The challenge is to evaluate and identify the most influential biomarkers among all the measured candidates. A spline-based nonparametric additive model with monotonic constraints is developed to flexibly model the biomarker dose-response relationships. We propose an innovative stepwise forward nonparametric variable selection approach to identify respondent biomarkers. Features of this method include flexibility of dose-response relationship, robustness result against correlated biomarkers, and ease of implementation while maintaining the important aspects of conventional variable selection approaches.

Quantifying the Usefulness of PD Biomarkers in Phase II Screening Trials of Oncology Drugs

✱Eric B. Holmgren, Genentech, Inc, 1 DNA Way, South San Francisco, CA 94306, *erich@gene.com*

Key Words: biomarker, surrogate, meta analysis, Phase 2

The increasing cost and duration of Phase II trials are leading drug and biotechnology companies to look for ways to speed up this part of the drug development process. One approach is to rely on pharmacodynamic markers to provide a quick assessment of a drug's activity. If such a marker is a correlate of survival and the marker is also sensitive to drug effects, then between group differences in changes in the marker may correlate with between group differences in survival. In this paper we ask, will the use of the PD marker in Phase II as a screen for drugs that will pass to Phase III improve the efficiency of the drug development process? This paper extends the framework presented in Holmgren [1] for analyzing the efficiency of drug development. Reference 1. Holmgren EB. Are Phase 2 screening trials in oncology obsolete? *Stat Med* 2008; 27 (4) 556-567 (DOI 10.1002/sim.2989)

Impact of Measurement Variability in PROs and How Under- or Over-Reporting of Extreme Valued Response May Influence Conclusions

✱Tammy Massie, U.S. Food and Drug Administration, 1401 Rockville Pike, Rockville, MD 20852, *Tammy.Massie@fda.hhs.gov*

Key Words: Patient Reported Outcome, Extreme Values, Categorical Response, Treatment Effects

Patient reported outcomes (PRO's) are included as endpoints in many comparative clinical studies. Within PROs, subjects may provide categorical responses over time. Data from each subject may be analyzed utilizing statistically valid categorical data analysis methodologies based on: change from baseline and difference between groups at preset time points. In cases of categorical responses, in which there exist few responses to choose from (ex. Likert scale with 5 responses) the over-use or hesitancy to use extreme values, may influence results and ultimately conclusions regarding treatment effect. One goal of this talk is to determine what, if any, impact potential under-reporting of extreme values may have on statistical analysis of categorical PROs. Simulations in SAS or S-Plus will be utilized to explore the potential influence of under-reporting of extreme values on conclusions.

199 Contributed Poster Presentations

Biometrics Section, Biopharmaceutical Section, Business and Economics Statistics Section, IMS, Section on Bayesian Statistical Science, Section on Government Statistics, Section on Health Policy Statistics, Section on Physical and Engineering Sciences, Section on Risk Analysis, Section on Statistical Computing, Section on Statistical Consulting, Section on Statistical Graphics, Section on Statistics and Marketing, Section on Statistics in Epidemiology, Social Statistics Section, WNAR, Section on Nonparametric Statistics

Monday, August 4, 2:00 p.m.–3:50 p.m.

Four-Week Repeated Dose Oral Pharmacology Study in the Beagle Dog

✱Amparo Y. Castro, Hasselt University, Agoralaan - Building D, Diepenbeek, International 3590 Belgium, *acastrosan@gmail.com*; John Maringwa, Hasselt University; Helena Geys, Johnson & Johnson Pharmaceutical R&D, LLC

Key Words: canine obesity, cholecystokinin, veterinary drug, mixed models, baseline correction

The main objective is to evaluate the effect of a Cholecystokinin (CCK) antagonist component jointly with an inhibitor on lowering the food consumption. Twenty four obese dogs were randomized to four treatments and the variables CCK and food consumption were recorded at different time points. To asses the effect of the treatments we use linear mixed models, including a baseline correction. For the response variables, we started from the most elaborate model, then we try to simplify it using the likelihood ratio test and the AICC. To check the model we use Local Influence Analysis and the comparison between the observed and the predicted values. We found statistically significant treatments differences, although non-constant through time. In some cases the baseline correction improved the model fitted but affects the significance of the variance components.

Bayesian Optimal Design for Nonlinear Combinations of Parameters in Gaddum/Schild Model

✱Zhongwen Tang, Kansas State University, 66502, *zhongwent@yahoo.com*

Key Words: Gaddum/Schild, Bayesian Optimal Design, EC50, IC50

Gaddum/Schild model has been used in a number of cases to model response to combined agonist and antagonist stimuli. In practice, researchers are most interested in estimating two parameters, EC50 and IC50. In our work, we developed a Bayesian optimal design method for estimating both EC50 and IC50. An R package is developed to implement the method.

Symbolic Data Analysis Applied to Census Data

✱Antonio Giusti, University of Florence, Viale Morgagni, 59, Department of Statistics, Firenze, I 50134 Italy, *giusti@ds.unifi.it*; Laura Grassini, University of Florence

Key Words: Census Data, Local Labor Systems, Symbolic Data Analysis, Agriculture, Statistical Information System

The paper analyses the structural characteristics of agriculture activities in the Italian Local Labor Systems (LLS: a partition of the Italian territory, provided by ISTAT), on the basis of the microdata collected during the last Italian General Agricultural Census. The main aim is to provide a picture of the economic activities of the LLSs by integrating Agricultural Data with the Industrial Census Data. The Italian economy is characterized by small and medium sized firms, whereas large companies are very few. Hence, the features of the Italian economy cannot be properly understood without taking into account the territorial organization. The main characteristic of the paper is the use of the symbolic data analysis approach to aggregate the LLSs into non overlapping clusters.

An Adaptive, Two-Stage, Dose Exploration Design for the Estimation of a Human Colonizing Dose 50 (HCD50) and Human Colonizing Dose 90 (HCD90)

✱Yu-Hui H. Chang, The University of Iowa, 200 Hawkins Drive C22-GH, Department of Biostatistics, Iowa City, IA 52242, *yu-hui-huang@uiowa.edu*; Kathryn Chaloner, The University of Iowa; Patricia Winokur, The University of Iowa; Michael Apicella, The University of Iowa; Wei Zhang, Boehringer Ingelheim Pharmaceuticals, Inc.

Key Words: Adaptive design, Bayesian design, sequential design

Healthy volunteers were inoculated with nontypeable Haemophilus influenzae with the ultimate goal of estimating the doses at which 50% and 90% of subjects become colonized (the HCD50 and HCD90). The initial study was designed to guide in the design of subsequent studies and to explore the dose response relationship. The fifteen-subject study was designed in two stages, with the first six subjects allocated sequentially. The design was chosen based on heuristic arguments. This design and a number of alternative designs are evaluated by simulation, under both Bayesian and frequentist criteria. The alternative designs are Bayesian and use both myopic and backward induction methods. The Bayesian designs provide better estimates. Results from the original study are used to design a subsequent study to better estimate the HCD50 and HCD90 using Bayesian methods.

Identification of Penicillium Species Using Discriminant Analysis

✱James Slaven, National Institute for Occupational Safety and Health, 1095 Willowdale Rd, MS 4050, Morgantown, WV 26505, *cto8@cdc.gov*; Justin Hettick, National Institute for Occupational Safety and Health; Brett Green, Sporometrics; Amanda Buskirk, National Institute for Occupational Safety and Health; Don Beezhold, National Institute for Occupational Safety and Health; Michael Kashon, National Institute for Occupational Safety and Health

Key Words: Discriminant, Cluster, MALDI-TOF MS

Identification of morphological colony features of fungi can be subjective. A method of classifying Penicillium species using MALDI-TOF MS has been established to improve species identification. Validation of this method was established with discriminant and cluster analyses. Stepwise analysis identified the significant peaks generated by MALDI-TOF MS. These peaks were used to determine classification error rates between species. Canonical analysis was used to generate linear combinations of the original data set to reduce its dimensionality. Stepwise analysis kept 19 of the 42 peaks, which were classified by strain with discriminant analysis with a 0% error rate. Two canonical coefficients were sufficient to explain over 95% of the data variability, which are shown graphically. Overall, discriminant analysis methods work very well in differentiating Penicillium species.

Cluster Analysis and Predicting Transcription Factor Regulatory Network

✱Dongseok Choi, Oregon Health & Science University, 3181 Sam Jackson Park Road CB669, Portland, OR 97239, *choid@ohsu.edu*; Zhixin Kang, University of Illinois at Chicago; George C. Tiao, The University of Chicago

Key Words: transcription factor, motif, module, microarray, cluster, gene expression

Microarray technology enables us to explore gene expression patterns in different diseases or tissue types. Transcription factor regulatory network does a crucial role in regulating gene expression patterns. However, there are often too many genes expressed differentially, which makes prediction of transcription factor regulatory network challenging due to too much information or noise. In this presentation, we will explore various clustering methods to investigate how they can be used to overcome the challenges.

Estimating Weekly Stocks of Other Oils

✱Ruey-Pyng Lu, Energy Information Administration, EI-70, 1000 Independence Avenue, SW, Washington, DC 20585, *ruey-pyng.lu@eia.doe.gov*

Key Words: Multiple regression model, Petroleum Supply Monthly, Unobserved components model, Weekly petroleum Status Report System

The Energy Information Administration collects data on stocks of other oils on a monthly basis. However, weekly data are needed to obtain the estimates of total petroleum stocks and petroleum products supplied that appear in the Weekly Petroleum Status Report (WPSRS). We evaluated several methods to develop weekly estimates by comparing the “monthly from weekly” estimates we obtained from these methods to Petroleum Supply Monthly (PSM) data. The statistical methods under investigation for estimating the WPSRS “other oils stocks” included multiple regression models and unobserved components models (UCM). The UCM model using monthly data showed the most plausible Mean Absolute Percent Error and the smallest error range. It is both a replicable method and a method that produced estimates that corresponded closely with the PSM data.

Autopsy Status of Homicide Victims and Medical Examiner/Coroner System in the United States in 1979–1994

✱Hsiang-Ching Kung, National Center for Health Statistics, 3311 Toledo Rd., Hyattsville, MD 20782, *hck0@cdc.gov*; Xinhua Liu, Columbia University; Donna L. Hoyert, National Center for Health Statistics

Key Words: time trend, hierarchical bayesian linear model, autopsy, homicide, death certificate, medical examiner/coroner system

The purpose of this study was to examine patterns of the annual autopsy rate on homicides by the medical examiner/coroner (me/c) system overtime 1979–1994, and whether the revision of death certificate in 1989 had effect on the reporting autopsy rate. Study variables selected from death certificates in 1979–1994. We first calculated annual autopsy rate with standard errors for each me/c system, then used hierarchical Bayesian linear models to examine the temporal trend in the proportation and test whether there were pre-post 1989 changes in the annual autopsy rate. The coroner system had increasing time trend in annual autopsy rates from 87.6% in 1979 to 93.9% in 1989 and jump to 96% at 1990 and 1991 and kept higher to 97.2%. The medical examiner system and mixed system had similar pattern in the annual autopsy rates, from 94% in 1979 to 97% in 1989, then relatively stable.

Comparisons of the Uses of Logistic and Probit Regression in the Aerospace and Biomedical Fields

✱Peter Hovey, University of Dayton, 300 College Park Ave, Dayton, OH 45469-2316, *Peter.Hovey@notes.udayton.edu*; Rafe Donahue, Vanderbilt University

Key Words: logistic, probit, engineering, biomedical, inspections

The tools of logistic and probit regression have traditionally been used in biomedical research to model the probability of an outcome such as death or recovery as a function of the amount of exposure to a toxin, the dosage of a medication, or other continuous predictors. Recently, the aerospace industry has used logistic and probit regression to model the probability of detecting cracks as a function of the size of the cracks. Several examples of how these tools are used in the biomedical and aerospace fields will be presented and contrasted and compared.

Equivalency Criteria in Pharmaceutical vs. Engineering Applications

✱Elizabeth Whalen, The Boeing Company, P.O. Box 3707 MC 7L-22, Seattle, WA 98124, *elizabeth.a.whalen@boeing.com*

Key Words: hypothesis tests, power, type I error

Equivalency criteria in pharmaceutical industry formulate the hypothesis to demonstrate equivalency as data deviate from the null hypothesis. In engineering applications, the traditional hypothesis setting is used so that failure to reject the null hypothesis is deemed to conclude equivalency. The limitations and strengths of these approaches are helpful for practitioners to decide which approach is more appropriate for their applications. It is also important to note their differences so that industry standards or regulatory agencies would deal with the certification process under these methods accordingly. This poster presentation illustrates these approaches with some examples and discusses open issues for both methods.

Power Transformations Based on Data from the National Health and Nutrition Examination Surveys

✱Margaret D. Carroll, National Center for Health Statistics, 3311 Toledo Road Room 4413, Hyattsville, MD 20782, *mdc3@cdc.gov*

Key Words: skewed distributions, square root transformation

In analyzing data from the National Health and Nutrition Examination Surveys (NHANES) conducted by the National Center for Health Statistics/Centers for Disease Control and Prevention, highly skewed variables are often encountered, for example serum vitamin C and lipoprotein a. Consequently, a transformation is needed to construct confidence intervals and to test statistical hypotheses. Frequently, the log transformation is used. However, it does not always reduce skewness. Therefore, alternative transformations are needed. Here, examples including the square root and the fourth root will be applied to NHANES III and NHANES 1999–2006 data. Guidelines for determining whether a transformation is needed, finding the optimal transformation and applying it in confidence interval construction and hypotheses testing both in a univariate and multivariate context will be outlined.

Analysis of the Growth Rate of Prostate-Specific Antigen of Prostate Cancer Patients by a Linear Spline Model with Mixed Effects

✱Suyan Tian, Rockefeller University, 1230 York Avenue, New York, NY 10065, *stian@rockefeller.edu*

Key Words: Prostate cancer, Prostate specific antigen (PSA), Linear spline models, Random effects

A Phase I/II clinical trial to evaluate the immunogenicity of dendritic cell vaccine (DC/PC3) in 24 patients with prostate cancer was conducted at the Rockefeller University. The change in growth rate of prostate specific antigen (PSA) between pretreatment, treatment and post-treatment phases was measured using a linear spline model with two knots. Random effects were considered to account for the heterogeneous treatment effect and the repeated measures structure. The final model used was a mixed linear spline model with 3 random effects based on AIC/BIC and scientific plausibility. A barely statistically significant decrease in the slope of log (PSA) from treatment to post treatment phase (B=-0.061/month, p=0.045) suggests there might be a slowed disease progression after vaccination. The change between pretreatment and post-treatment phases was not significant.

Applications of Nonclinical Biostatistics in Biotechnology Industry

✱Priya Kulkarni, Genentech, Inc., 1 DNA Way, MS 441B, South San Francisco, CA 94080, *preeya@gmail.com*

Key Words: biotechnology, consulting, applications

Applications of biostatistics in biotechnology industry, sans clinical trials, span a wide range including basic research, assay optimization and validation, process optimization, process control, etc. A consistent challenge for a biostatistician working with scientists and engineers is to know their subject matter and use that knowledge to apply statistical methods creatively to solve practical problems. We describe some of the interesting statistical applications in biotechnology industry through various case studies.

Market Segmentation of Customers' Satisfaction and Their Willingness To Pay for the Cable TV Industry of the Chung-Hua Area in Taiwan

✱Nai-Hua Chen, Chienkuo Technology University, No. 1 che-shu north rd., Chung-Hua, 403 Taiwan, *nhc@cc.ctu.edu.tw*; Ching-Tong Wu, Chienkuo Technology University; Shu-jing Chiu, Chienkuo Technology University; Chung Tu, Chienkuo Technology University

Key Words: self-organizing maps, ordinal logistic regrssion, customer's satisfaction

The penetration rate in Taiwan is more than 70 percent. However, the development in digital technology, satellite TV and internet highly influences the market of the cable TV. The cable TV company needs a further plan in management strategies. In this paper, a neural network technique is used in market segmentation. Then the ordinal logistic regression is applied in estimating the willingness to pay for the monthly fee of the cable TV. Three clusters are named as heavy-viewer, loyal-viewer and young-viewer based on the variables of importance in Cable TV quality, demography and the viewing experience. Represents of the heavy-viewer spend more than six hours a day in watching TV. Their satisfaction for the cable TV quality is lowest but they are willing to pay highest in three groups. Loyal-viewers have been installed cable more than eight years with highest satisfaction in TV program.

A Natural Experiment Examining Proximity to Emergency Department (ED) as a Predictor of Frequent ED Use Among Asthmatic Children in the Inner City

✱Henry T. Bahnson, Rho, Inc., 5 Fall Circle, Durham, NC 27713, *teebahnson@yahoo.com*; Agustin Calatroni, Rho, Inc.; John Schwarz, Rho, Inc.; Herman Mitchell, Rho, Inc.

Key Words: natural experiment, quasi-experiment, emergency department, asthma, inference, confounding

Emergency Department (ED) use is high among inner city children with asthma. A novel approach is presented to examine the relationship between home proximity to EDs and rate of ED utilization in this population. While exploring this relationship in 937 children with asthma in 7 major urban areas, we discovered that nearly 1/3 of the population had moved during the 2-year study. This allows us a unique opportunity to examine this relationship in a natural experiment design and compare the differences in statistical inference to the originally conceptualized cross-sectional approach. GIS (Geographical Information System) technology was utilized to measure home proximity to ED. The inverse relationship of distance from residence to ED and rate of ED utilization is fit using GEE methodology. A discussion of confounding and statistical inference is emphasized in this quasi-experiment.

Cancer Mortality in Texas

✱Ferry Butar Butar, Sam Houston State University, 3657 Red Bud ln, Huntsville, TX 77340, *butar@shsu.edu*; Ananda Bandulasiri, Sam Houston State University

Key Words: SMR, empirical bayes, hierarchical bayes

We will use a hierarchical Bayes methodology to map the cancer mortality for each county in the state of Texas. We will use the data from year 1990–1997 as an example. We will map it using Poisson, normal and binomial models. We use a Monte Carlo simulation to find the measure uncertainty.

Undergraduate Retention Using Cox Regression

✱Barbara Warsavage, California State University, East Bay, 25800 Carlos Bee Blvd, Hayward, CA 94542, *barbara.warsavage@csueastbay.edu*

Key Words: Student Retention, Educational Research, Cox Regression, Higher Education

Undergraduate student retention at a California State University East Bay is a strong indicator of student success and of the health of the university. At California State University East Bay a large study of retention is being done by the Assistant Vice President for Institutional Research, Colin Ormsby, PhD. Retention is difficult to study because no one factor seems to be related. Warsavage (2007) showed that neither the quality of the high school or the student SAT score were clear determinants of retention. I will look at retention using a variety of methods including the Cox regression model. The purpose is to inform policy makers of those factors or combinations of factors that indicate when intervention may improve retention and student success.

An Algorithm for Estimating Power and Sample Size for Logistic Models with One or More Independent Variables of Interest

✱Jay Northern, University of Arkansas for Medical Sciences, AR 72205, *jay.northern@arkansas.gov*; D. Keith Williams, University of Arkansas for Medical Sciences; Zoran Bursac, University of Arkansas for Medical Sciences

Key Words: Power, Sample Size, Multiple Logistic Regression

Determination of power, and sample size in planning a study for logistic regression has been proven to be a relatively simple task when investigating a dependent variable (binary outcome) and one independent variable. Extending beyond one independent variable, that is, testing the significance for a set of independent variables presents a problem for the investigator in estimating power and sample size. Currently, power and sample size computer programs are arguably somewhat limited for logistic regression. We present an algorithm that estimates samples size for logistic regression with one or more independent variables while controlling for other covariates.

Effect on Power of Categorizing Continuous Variables

✱T. Robert Harris, The University of Texas School of Public Health, 6011 Harry Hines Blvd., v8.112, Dallas, TX 75390-9128, *TRobert.Harris@UTSouthwestern.edu*

Key Words: power, misspecification, categorical data

Theory leads to the expectation that hypothesis tests based on a correctly specified model will have greater power than alternatives based on a misspecified model. The extent of difference in power due to using available continuous variables or categorizing some or all of them is examined via simulation, in several situations. These include the case where the correct model is in fact linear dependence of outcomes on continuous predictors; the case in which the relationship is correctly modeled by a categorical predictor; and situations which fall between these poles. Additionally, linear and logistic regressions using continuous and categorized versions of the same dependent variable are analyzed. Comparisons depend on sample size, effect size, and correlations among independent variables. Differences in power are substantial, in some cases more than double.

An Organized Investigation of Sample Size in Independent and Paired Dichotomous Data

✱Kang-Hsien Fan, University of Louisville, 2400 Mellwood Ave. # 907, Louisville, KY 40206, *k0fan001@louisville.edu*; Linda J. Goldsmith, University of Louisville

Key Words: sample size, dichotomous data, paired data, statistical power, McNemar's test, matched data

Efficient research is desirable and is characterized by careful sample size calculation. In dichotomous data, researchers believe that paired data will decrease the sample size and reduce the cost of experiments. Formulas for sample size calculation for paired dichotomous data have been developed, but the methods have not been organized by increasing levels of dependence. In our study, we develop an organization method for sample size in independent and paired dichotomous data which indexes data tables according to increasing degree of agreement. Using the Bahadur r correlation for paired dichotomous data, we show that with higher correlation sample size does decrease. However, we note that the increased cost of matching and administering paired samples may diminish any savings due to smaller sample size.

Ordinal Classification Approach Using Bagged Classification Trees and the Proportional Odds Model as Splitting Criteria

✱Tobias Guennel, Virginia Commonwealth University, 730 East Broad Street, P.O. Box 980032, Richmond, VA 23298-0032, *tguennel@vcu.edu*

Key Words: Ordinal Classification, Proportional Odds Model, Bagging, Splitting Criteria, Cumulative Logits, Classification Trees

This work proposes a new approach to classification problems when the response variable is ordinal. The method makes use of classification trees (CTs) stabilized through bootstrap aggregating (bagged CTs). However, in place of the standard splitting criteria (i.e., Gini, entropy), the proposed method incorporates the ordinal structure into the splitting procedure by using proportional odds models to uncover dependencies between the ordinal response and the explanatory variables. The performance of bagged CTs using the modified splitting rules are compared to ordinary bagged CTs by reporting misclassification rates using one simulated data set and two real data sets. The results show significantly reduced misclassification rates for the simulated data set and slightly reduced misclassification rates for the real data sets, which suggests further research in this area.

Efficient Selection of Prototypes Using Expert Ratings

✱William E. Miller, Centers for Disease Control and Prevention, 1095 Willowdale Road, Morgantown, WV 26505-2888, *wem0@cdc.gov*

Key Words: Monte Carlo simulation, Methodology, Allocation, Ordinal classification, Random block design, Data coding

Monte Carlo simulation was used to compare methods for using expert ratings to choose objects or subjects that are prototypical of various conditions. These simulations were done in preparation for selecting x-ray films, for training purposes, which are representative of specified categories of ordinal classifications using ratings given by chest physicians. The probability of correct allocation or assignment to a specified category may depend on the statistic or metric used for comparison, the choice of weights for the raters, and the coding of the data. Various methods were compared using a variety of multinomial probabilities for classifications with two, three, and five categories. A generally effective approach uses the mean-absolute error together with a composite weight which combines latent class and mean-squared error estimation.

Assessing Regression Modeling with Ordinal Responses

✱Kao-Tai Tsai, BMS, 34 Baldwin Dr., Berkeley Heights, NJ 07922, *tsai0123@yahoo.com*

Key Words: Ordinal data, goodness of fit

In clinical trials, data collected are very often categorical ordinal data. These can include the binary yes or no answers or categorical answers indicating the severity of disease symptoms. Logistic or Poisson regression is the common method used to analyze these kinds of data. However, given the discrete nature, goodness of fit by examining the residuals from regression is not as useful as in the case for the continuous variables. Landwehr, et al. and Cook had proposed methods to assess the goodness of fit for the binary data. In this paper, we will extend their results to general categorical responses. We will assess the goodness of fit via both smoothing and the residuals. Data from a recent clinical trial will be used to illustrate the proposed method.

Exact Inference for Ordered and Nominal Categorical Data

✱John D.S. Hwang, B.R.S.I., East Hanover, NJ 07936, *dshwang@optonline.net*; James S. Lee, Daiichi Sankyo Pharma Development

Key Words: Exact Inference, Categorical Data, Multinomial, Conditional Distribution, Interaction

Clinical trials often involve rating scale or nominal data. The underlying distribution is multinomial. Exact inference for comparison of k >= 2 treatments has been developed. Consider now interaction for nominal data in a 2-treatment, k-level, 2-way layout. Parameters representing interaction are identified. Existence and derivation of a conditional distribution which depends on these parameters are given, and may be employed for statistical testing and interval estimation. For ordered categorical data, first analyze them as nominal data as above. Then convert the observed ordinal data to Wilcoxon or Kruskal-Wallis rank statistic. Probability associated with the rank statistic is related to those associated with the corresponding nominal data. When there is no interaction, similar model-based exact analysis is given for the "common" treatment effect under null and non-null hypothesis.

Development and Internal Validation of the Prediction Model To Identify Thrombotic Thrombocytopenic Purpura (TTP) Patients at Risk for Relapse

✱Xiaoning Li, University of Oklahoma Health Science Center, 402-A Wadsack Dr., Norman, OK 73072, *xiaoning-li@ouhsc.edu*

Key Words: Internal validation, logistic regression, prediction model

Despite most TTP patients respond to plasma exchange (PE), a significant number of patients relapse. The goal was to develop and validate a clinical prediction model to identify TTP patient at risk for relapse using Oklahoma TTP-HUS Registry data. Complete data (253 patients) are available for all persons with clinically suspected TTP-HUS in Oklahoma (1/1/1989 - 01/31/08). Patients survived at least a month were randomly divided into development (n=69) and validation (n=69) group. Multivariate logistic regression analysis was used to derive a best fitting model predicting relapse from presenting characteristics based on the development group data. Then, the best fitting model was validated using the validation group data.

The Partition Logistic Regression Model

✱Ying Liu, Virginia Polytechnic Institute and State University, 2400 Foxhunt Ln, Apt L, Blacksburg, VA 24060, *yingliu24@yahoo.com*; Shi-shien Yang, Kansas State University

Key Words: logistic regression, lack-of-fit

To improve the power of the lack-of-fit tests for logistic regression model, we present a partition logistic regression model, which is powerful to detect the overall lack-of-fit, between groups lack-of-fit and within groups lack-of-fit. It is recommended that the overall lack-of-fit test should be performed first. If the null hypothesis is not rejected, we simply conclude that lack of fit for the assumed model can't be detected; otherwise, the between and within group lack-of-fit test should be carried out to figure out the type of lack of fit. The partition model also can be used to detect nature of lack-of-fit in the tails or middle part. The proposed model is flexible according to specific part of interest.

College Desirability: A Multivariate Statistical Analysis

✱Lindsay Moomaw, SUMSRI, *lmoomaw@bw.edu*; Andrea Austin, SUMSRI; Terrell Felder, SUMSRI

The colleges and universities across the United States are all unique. To quantify how institutions of all sizes measure up, multivariate techniques of Principal Component Analysis, Factor Analysis, and Discriminant Analysis are used fittingly and effectively, producing a valid, unbiased evaluation of each school, and also a model to gauge any chosen seminary. The method of Principal Components reduces the number of variables, focusing on those with efficacy while Factor Analysis provides a data reduction to explain the variability of the college or university statistics. Finally, a Discriminant Analysis of the data classifies the schools and establishes a method of accurate prediction.

Discovering Sparse Covariance Structures with the Isomap

✱Amy Wagaman, The University of Michigan, *awagaman@umich.edu*

Key Words: covariance estimation, sparsity, regularization, manifold projections

Regularization of covariance matrices in high dimensions is usually either based on a known ordering of variables or ignores the ordering entirely. This paper proposes a method for discovering meaningful orderings of variables based on their correlations using the isomap, a nonlinear dimension reduction technique designed for manifold embeddings. These orderings are then used to construct a sparse covariance estimator, which is block-diagonal and/or banded. Finding an ordering to which banding can be applied is desirable because banded estimators have been shown to be consistent in high dimensions. We show that in situations where the variables do have such a structure, the isomap does very well at discovering it, and the resulting regularized estimator performs better for covariance estimation than other regularization methods that ignore variable order, such as thresholding.

Application of Transportation Problem to Prediction Model in Orthodontics

✱Olena Tsvirkunova, Align Technology Inc., 831 Martin Av., Santa Clara, CA 95050, *otsvirkunova@aligntech.com*; Michael Zakharevich, Align Technology Inc.; Vadim Matov, Align Technology Inc.

Key Words: Prediction, Transportation Problem, Parameter Dependencies, Penalty Function

We develop Prediction Model for Success of Orthodontic Treatment with Invisalign product. We start with a set of categorical variables describing the diagnostic of initial dentition and we build a model to predict success/failure. Standard approach assumes unrealistic conditions (independence of variables) which are difficult to expect in practice. Another complication is the skewed distribution of training data. (It includes only patients that had chosen Invisalign treatment). To compensate, we consider a modification of the standard approach taking into account heuristic knowledge from experts. We express knowledge in the form of parameter dependency graph and penalty functions for unrealistic relation between parameter values. Adding expert knowledge reduces the model to Transportation Problem of Linear Programming. We propose an iterative method providing quasi-optimal solution.

Neural Network Monitoring of Poisson Data

✱Benjamin M. Adams, The University of Alabama, Box 870226, Tuscaloosa, AL 35487, *badams@cba.ua.edu*; Yousef Al-Hammadi, The University of Alabama

Key Words: control charts, c-charts, statistical process control, quality control, attribute data

The performance of neural network monitoring schemes for Poisson data are compared with more traditional methods such as the C-chart. Desirable neural network architectures and their performances are reported. Bounds for minimal training data requirements for good ARL performances are reviewed for neural network schemes. Finally, the robustness of the various monitoring schemes to the Poisson assumption is reported.

S-Plus and R Package for Least Angle Regression

✱Tatiana Maravina, University of Washington, Department of Statistics, Box 354322, Seattle, WA 98195-4322, *maravina@u.washington.edu*; Tim Hesterberg, Google, Inc; Chris Fraley, Insightful Corporation

Key Words: regression, regularization, L1 penalty, variable selection, large datasets

Least Angle Regression is a promising new technique for variable selection applications, offering a nice alternative to stepwise regression. It provides an explanation for the similar behavior of Lasso (L1-penalized regression) and forward stagewise regression, and provides a fast implementation of both. We'll demonstrate an open-source S-PLUS/R package "glars" for generalized least angle regression, extending the work by outside collaborators. In particular, we'll present the scalable lars (a prototype S-PLUS package "sclars"), an extension of LAR and LASSO to the datasets which can't be stored in the memory due to the very large number of observations. See *www.insightful.com/Hesterberg/glars*.

Support Vector Machines with Recursive Feature Elimination for the Creation of a Diagnostic Tool

✱David H. Henderson, Insightful Corporation, 1700 Westlake Avenue, North, Suite 500, Seattle, WA 98109, *dnadave@insightful.com*

Support vector machines with recursive feature elimination was used to create a classifier for exposure and dosage of exposure to the toxin ANIT in 22 rat samples. Genomic and binned mass spectrophotemetric data were used as inputs to the SVM-RFE procedure producing a classifier with 32 features total from the combined sources. This classifier had a 2% false positive rate and a 9% false negative rate estimated using a 10 fold cross validation. This research was funded from a US Air Force SBIR grant.

On Nigerian Statistical Capacity Building as a Panacea for Achieving Millennium Development Goals (MDGs)

✱Dallah Hamadu, University of Lagos, Department of Actuarial Science and Insurance, Lagos, 9600001 Nigeria, *dallaram2007@yahoo.com*; Ray Okafor, University of Lagos

Key Words: Critique of Nigerian Statistical System, Statistical Capacity Building, MDGs, PARIS21, Addis Ababa Plan for Action (AAPA), NBS

The design and implementation of National Strategies for the Development of Statistics (NSDS) in developing countries is critical in achieving MDGs. Development planning in Nigeria which began in the 1960s after independence experienced difficulties due to lack of accurate and reliable data. Nevertheless, Nigeria being the most populous country in Africa has been seriously involved in most novel initiatives for building statistical capacity in Africa (AAPA, PARIS21). This paper critically assesses the Nigerian Statistical System and explores factors militating against the implementation of newly initiated NSDS, which is crucial to meet the government well publicized Vision 2015 MDGs strategies. Moreover, several examples from National Bureau of Statistics (NBS) and Central Bank of Nigeria (CBN) are provided to highlight the severity of inadequacy of existing system.

Graphical Displays of Safety Performance Functions

✱Barbara A. Bailey, San Diego State University, Department of Mathematics and Statistics, 5500 Campanile Drive, San Diego, CA 92182, *babailey@sciences.sdsu.edu*

Key Words: neural network, deriviatives

Safety Performance Functions (SPF) are accident prediction models that relate traffic exposure, measured in Annual Average Daily Traffic (AADT) to safety, measured in the annual number of accidents per mile (accidents/mile per year). A flexible nonlinear neural network model is used to fit the data and describe the underlying relationship between safety and exposure data. Estimates of the derivative of the functions and comparison of the distribution of derivatives can give insight into the SPF. On uncongested segments the number of crashes increases only moderately with increase in traffic; however, once some critical traffic density is reached, the number of crashes begins to increase at a much faster rate with increase in traffic.

200 ASA President's Invited Address

ASA, ENAR, IMS, SSC, WNAR

Monday, August 4, 4:00 p.m.–5:50 p.m.

Health Care Considerations for the Millennium

✱Mark B. McClellan, American Enterprise Institute, 1150 Seventeenth Street, NW, Washington, DC 20036, *Mark.McClellan@AEI.org*

201 Section on Statistical Education Roundtable with Coffee (fee event)

Section on Statistical Education

Tuesday, August 5, 7:00 a.m.–8:15 a.m.

Not Significantly Different: Nonstatisticians Who Teach College-Level Statistics

✱Patricia Rutledge, Allegheny College, 520 North Main Street, Box 39, Meadville, PA 16335, *patricia.rutledge@allegheny.edu*

Key Words: college-level teaching, non-statisticians

Many of us who teach college-level statistics courses across the disciplines are, in fact, not trained as statisticians. Although we possess a passion for statistics, our statistical education is less complete than that of our credentialed statistician colleagues. Given that the students in our courses, in all likelihood, are receiving the only formal statistics education they will encounter in college, it is clear the knowledge, skills, and abilities of those of us who are nonstatisticians teaching statistics have far-reaching effects. In this roundtable, we will meet for coffee and a discussion of issues of importance to nonstatisticians who teach statistics. We will seek to enhance teaching in college-level statistics courses, particularly those taught by nonstatisticians. Both nonstatisticians and statisticians are welcome.

202 Section on Statistics and the Environment Roundtables with Coffee (fee event)

Section on Statistics and the Environment

Tuesday, August 5, 7:00 a.m.–8:15 a.m.

Statistics and the Intergovernmental Panel on Climate Change

✱Peter Bloomfield, North Carolina State University, Department of Statistics, 2501 Founders Drive, Raleigh, NC 27695-8203, *bloomfield@stat.ncsu.edu*

Key Words: climate change, IPCC

The Intergovernmental Panel on Climate Change (IPCC) has issued four reports on the state of the science, most recently in 2007. Several statisticians have been directly involved in preparing those reports; others have collaborated in establishing the scientific basis for them. This discussion will focus on the role of statisticians in the IPCC process.

What Would a Statistician Do with a Climate Model?

✱Doug Nychka, National Center of Atmospheric Research, 1850 Table Mesa, Boulder, CO 80307-3000, *nychka@ucar.edu*

Key Words: climate, computer models, spatial statistics

What is a climate model? How is it used to understand potential changes in climate? Why should a statistician care? This is an informal discussion about some of the largest numerical models used by a scientific community and why concepts of uncertainty and model validation are central to making informed decisions about responding to global warming.

203 Section on Statistics in Epidemiology Roundtable with Coffee (fee event)

Section on Statistics in Epidemiology

Tuesday, August 5, 7:00 a.m.–8:15 a.m.

Statistical Analysis of Causal Intermediate Effects

✱Giovanni Filardo, Baylor Research Institute, 8080 N. Central Expwy, Suite 500, Dallas, TX 75206, *giovanfi@baylorhealth.edu*; Cody Hamilton, Edwards Lifesciences

Key Words: causal intermediate effects, confounding, mediator

Traditional modeling techniques assume an exposure of interest does not act as a mediator for other factors. For example, using standard regression techniques to assess the effect HIV-related knowledge (measured by standardized test) on a given behavioral outcome after adjusting for employment status and/or education would require that these factors only impact the outcome directly, and not through knowledge. Any effect that employment/education might have on behavior through knowledge cannot be estimated and/or accounted for. Factors such as employment/education may also have an effect on behavior that is not exclusively mediated through knowledge. In this case, knowledge should be considered as a causal intermediate factor. We will discuss accounting for confounding between exposure and other factors, allowing the exposure to function as a causal intermediate factor.

204 Section on Health Policy Statistics Roundtable with Coffee (fee event)

Section on Health Policy Statistics

Tuesday, August 5, 7:00 a.m.–8:15 a.m.

Prescription Drugs: Unsung Heroes of Disability

✱Michele J. Connolly, Sweetgrass Consulting LLC, 9556 Highwind Court, Columbia, MD 21045, *michelebabb@verizon.net*

Key Words: presciption drugs, disability, mental illness, health, health insurance, medical costs

Prescription drugs—like wheelchairs, scooters, and Braille—are an integral part of treating disabling conditions, maintaining function, and improving the lives of many people with disabilities. Examples include drugs for hypertension, asthma, cancer, diabetes, and mental illness. Yet, while we measure the effects of other assistive devices in population surveys and studies, we tend to ignore prescription drugs. If data on prescription drugs are collected, it is difficult to know whether these drugs were prescribed to treat acute conditions (e.g., antibiotics) or to maintain or improve function for a disabling condition (e.g., psychotropic drugs). Besides information about prescription drugs, we also need to measure a person's ability to routinely take drugs as prescribed, out-of-pocket costs, and insurance coverage. This roundtable will serve as a forum to discuss these issues.

205 Section on Survey Research Methods Roundtable with Coffee (fee event)

Section on Survey Research Methods

Tuesday, August 5, 7:00 a.m.–8:15 a.m.

Preventing Falsified Survey Data

✱Steven Pedlow, National Opinion Research Center, 55 E. Monroe, Suite 2000, Chicago, IL 60603, *pedlow-steven@norc.uchicago.edu*

Key Words: Curbstoning, Validation, Survey Fabrication

Despite the best efforts to recruit and train the best interviewers, there is still the risk interviewers will try to falsify cases. The standard technique of call-back validation is slow and unreliable in detecting falsified interviews. Various techniques have been developed, such as Benford's Law, sampling of audio files of interviews, looking for unlikely response combinations, and newer techniques such as GPS tracking. All these methods have limitations. The goal of this roundtable is to bring together researchers who can share their experiences to move toward the goal of preventing the falsification of interview data.

206 Section on Teaching Statistics in the Health Sciences Roundtable with Coffee (fee event)

Section on Teaching Statistics in the Health Sciences

Tuesday, August 5, 7:00 a.m.–8:15 a.m.

A New Statistics Student: The Translational Researcher

✱Carol Bigelow, University of Massachusetts Amherst, Program in Biostatistics & Epidemiology, Arnold House, Amherst, MA 01003, *cbigelow@schoolph.umass.edu*; Penelope Pekow, University of Massachusetts Amherst

Key Words: translational research, clinical epidemiology, basic statistics

The "bench-to-bedside" translational researcher is often a PhD trained in the basic sciences, a world steeped in the experimental method. Here, "like" individuals are randomized to study conditions set by the investigator and control of confounding is presumed by randomization. Discovery of effect modification is, at most, an exploration of biological synergism defined by study design. Appropriately, statistical training emphasizes analysis of variance, dose-response curve fitting, and the modeling of response surfaces, more generally. "Bedside" research, in contrast, often necessitates the use of observational study protocols and the measurement of dissimilar individuals. Statistical training of translational researchers might benefit from additional pedagogy in conditional likelihoods, epidemiologic study designs, control of confounding, and discovery of effect modification.

207 Introductory Overview Lecture: Harnessing Bibliographic Data

ASA, ENAR, IMS, SSC, WNAR, Current Index to Statistics

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Rejuvenating the Current Index to Statistics

✱Hadley Wickham, Iowa State University, Ames, IA 50010, *h.wickham@gmail.com*

The Current Index to Statistics indexes publications in statistics, probability and related fields and currently contains over 11,000 books and 250,000 journal articles and conference proceedings. How can we use modern web technologies to take advantage of this resource? How can we make it easy to find articles about a given topic, explore the body of work by an author, or locate interesting journal issues? In this talk, I will introduce a new web interface that adds navigation and sophisticated searching. I will discuss some of the challenges and opportunities that face CIS, and show how you can contribute to making this resource more complete and accurate.

Best Integrated Bibliographic Services

✱James Pitman, University of California, Berkeley, Statistics Department, Evans Hall #3860, Berkeley, CA 94720-3860, *pitman@stat.berkeley.edu*

While much bibliographic data is still distributed through subscription services, some large corporations are now offering free bibliographic search. There are opportunities in this environment for scholarly communities to signal their existence by selection, organization and statistical analysis of bibliographic data. The combination of signals from a network of such communities should then feed back to major open search engines and arenas like Wikipedia to provide a richly structured, navigable and reliable guide to human knowledge, open to anyone with an internet connection. To construct such a cheap and efficient network of professional information involves many problems of interest to statisticians and computer scientists: development of good computer (and human) interfaces to aid bibliographic data extraction, visualization, name disambiguation, document classification, and generation and matching of subject ontologies.

208 ASA College Stat Bowl I

ASA, ENAR, IMS, SSC, WNAR

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

There are 16 players in the tournament. There will be four first round games, each game consisting of four players. The placement of the players in the games was random, with the restriction of no one being in the same game as a person from the same university.

Game 1

✱Junheng Ma, Case Western Reserve University, 10900 Euclid Ave., Dept. of Statistics, Cleveland, OH 44106, *jxm216@case.edu*; ✱Thayasivam Umashanger, The University of Georgia, Statistics Department, 210 Rogers Road, Apt Q 204, Athens, GA 30605, *rtshanger@yahoo.com*; ✱Christopher

J. Rigdon, Arizona State University, 944 S. Terrace Rd. #204, Tempe, AZ 85281, *crigdon@asu.edu*; ✱Jian Zou, University of Connecticut, Statistics Department, 215 Glenbrook Road, U-4120, Storrs, CT 06269, *jian.zou@uconn.edu*

Game 2

✱John Garza, The University of Texas at San Antonio, Demography and Organizational Studies Department, 4980 USAA Blvd #1228, San Antonio, TX 78240, *lhr534@my.utsa.edu*; ✱Susan Buchman, Carnegie Mellon University, 5000 Forbes Avenue, Baker Hall 132, Pittsburgh, PA 15213, *sbuchman@stat.cmu.edu*; ✱Hongyuan Cao, The University of North Carolina at Chapel Hill, 881 Martin Luther King Jr. Blvd Apt. 18H, Chapel Hill, NC 27514, *hycao@email.unc.edu*; ✱Samsiddhi Bhattacharjee, University of Pittsburgh, PA 15232, *samsiddhi@gmail.com*

Game 3

✱Dan Polhamus, The University of Texas at San Antonio, Management Science and Statistics Department, 7721 Hunter Oaks, San Antonio, TX 78233, *Daniel.polhamus@utsa.edu*; ✱Liang Hong, Indiana University Purdue University, Indianapolis, Mathematical Statistics Department, 5905 Coquina Key Drive, Apt G, Indianapolis, IN 46224, *lihong@iupui.edu*; ✱Adam Molnar, Bellarmine University, 2001 Newburg Road, Louisville, KY 40205, *amolnar@bellarmine.edu*; ✱TBA

Game 4

✱Jing Xu, The University of Georgia, Statistics Department, 300 Rogers Road, Apt R212, Athens, GA 30605, *jasperxj@uga.edu*; ✱Brian Hartman, Texas A&M University, Statistics Department, 3625 Wellborn Rd. #308, Bryan, TX 77801, *bhartman@stat.tamu.edu*; ✱Anne M. Hansen, University of California Riverside, Department of Statistics, Riverside, CA, 92507, a*hans004@ucr.edu*; ✱TBA

209 Innovative and Controversial Approaches to Student Assessment

Section on Statistical Education
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Dirk Gently's Guide to Holistic Assessment

✱Andrew S. Zieffler, The University of Minnesota, 167 Educational Sciences Building, 56 East River Road, Minneapolis, MN 55455, *zief0002@umn.edu*

Key Words: Assessment, Statistics Education, Cooperative Learning, Curriculum Evaluation, Psychometrics, Teaching

In the books by Douglas Adams, Dirk Gently labels himself as a "holistic detective" who makes use of "the fundamental interconnectedness of all things" to solve the whole crime, and find the whole person (Adams, 1987). This paper will focus on several different aspects of assessment that have been implemented in multiple sections of an introductory statistics course at the University of Minnesota to create a more holistic course. These include cooperative (group) assessment, and the use of psychometric methods to evaluate and improve both the assessments and curriculum.

Grading Statistics Like Statisticians

✱Daniel Kaplan, Macalester College, Math/CS, 1600 Grand Avenue, Saint Paul, MN 55105, *kaplan@macalester.edu*

Key Words: Assessment, grading, modeling

Scores from exams and problem are usually entered into a spreadsheet for computation of the final grade as a weighted mean. Of course, the student's performance is only one factor that shapes the score; the quality of the item and the instruction also shape the score. To adjust for these, I treat the calculation of a grade as a modeling problem. For each item, the score is modeled by a categorical variable with one level for each student and another variable with one level identifying the item. This makes it possible to grade not just the student, but the item itself. Influential points and outliers can be identified and missing data handled in a nonpunitive way. I couple this modeling approach with an on-line assignment collection system that provides a high level of modularity so that the reliability and informativeness of individual questions can be assessed.

Against Fairness

✱George Cobb, Mt. Holyoke College, Department of Mathematics and Statistics, 404C Clapp Laboratory, South Hadley, MA 01075, *gcobb@mtholyoke.edu*

Key Words: assessment, statistics education, fairness, grade inflation

Natural variability ensures that all of us who think and care about our own teaching experiences become outliers in at least some ways. Statistics education has been enriched by our willingness to regard all points within the convex hull of these outliers as possible models for learning. In my talk, I'll describe some things I've learned about assessment from teaching in an environment that I've come to appreciate as highly unusual. With caveats, I shall argue against fairness, in favor of grade inflation, and for devaluing sincerity in favor of authenticity.

210 Statistical Challenges in Large-Scale Genetic and Genomic Studies ●▲

International Indian Statistical Association, Biopharmaceutical Section, WNAR, *CHANCE*
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Are a Set of Microarrays Independent of Each Other?

✱Bradley Efron, Stanford University, Department of Statistics, Sequoia Hall, 390 Serra Mall, Stanford, CA 94305-4065, *brad@stat.Stanford.EDU*

Key Words: independence test, correlated genes, microarrays

A typical microarray study might involve two groups of subjects: controls and treatments. Each subject provides material for his or her individual microarray, reporting some large number of genetic expressions at the same time, yielding an m by n data matrix "X," with m genes and n subjects. We expect the measurements down any one column to be correlated, since genes act in concert, making the rows of X correlated. However the columns, that is the microarrays, are usually assumed to be independent, an essential assumption for t-tests, cross-validation, permutation tests, etc. This talk concerns testing a matrix X for column-wise independence when the rows may be highly correlated. The effect of row-wise correlation is to reduce effective sample size—from 20426 to 17 in my main example.

Identifying Interactions and Model-Building in Genome-Wide Association Studies

✱Charles Kooperberg, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N / M3-A410, Seattle, WA 98109, *clk@fhcrc.org*

Key Words: GWAs, predictive models, interactions, epistasis

Many genome-wide association studies, in which 100,000s of SNPs are genotyped in studies with a few thousand participants are now being carried out. Some initial successes have been reported. Most of these results have come from single SNP analyses. It is generally assumed that many genes (SNPs) contribute jointly, either additively or with epistatic effects, to the risk if disease. Clearly because of the size of the problem it is harder to identify epistatic effects. In this talk we will discuss approaches to increase the power to identify such effects, as well as methods to combine significant single SNPs and identified epistatic effects to form genetic risk models.

Powerful Discovery of Genetic Associations in the Presence of Gene-Gene and Gene-Environment Interactions

✱Nilanjan Chatterjee, National Cancer Institute, 6120 Executive Blvd, Division of Cancer Epidemiology and Genetics, Rockville, MD 20852, *chattern@mail.nih.gov*; Julia Ciampa, National Cancer Institute; Idan Menashe, National Cancer Institute; Sheng Luo, Johns Hopkins University

Key Words: case-control study, genome-wide scan, score-test, pathway, multi-locus test

The completion of the Human Genome Project and rapid advancement of genotyping technologies now hold great promise for discovering the inherited causes of complex diseases by studying genetic variations across candidate genes, biochemical pathways and the whole genome. This talk will examine efficient strategies for "conditional search" for disease susceptibility loci using omnibus test statistics that can simultaneously account for genetic main effects and gene-gene/gene-environment interactions. The venerable and parsimonious Tukey's "one degree-of-freedom" model for interaction will be used in this context to develop highly powerful but simple score test of genetic association that can be implemented using standard software. Applications of the methods for the discovery of individual susceptibility loci as well as of global pathway effect will be demonstrated using real studies.

Efficient Statistical Analysis of Case-Control, Genome-Wide Association Study Data

✱Michael Boehnke, The University of Michigan, Department of Biostatistics, School of Public Health, Ann Arbor, MI 48109, *boehnke@umich.edu*

Key Words: genetics, association, stratification, winner's curse, type 2 diabetes

With the availability of a catalog of common human variants and the rapid drop in genotyping costs, genome-wide association studies now provide a powerful means to localize genetic variants that predispose to human diseases. Many such studies are now in progress. In this talk, I will focus on solutions to two statistical problems posed by genome-wide association studies: avoiding spurious results owing to population stratification by matching based on genome-wide marker data and correcting for overestimation of genetic effect size owing to the "winner's curse" by conditioning on the observation of a significant result. I will illustrate the value of these methods with data from the Finland-United States Investigation of NIDDM Genetics (FUSION) study.

211 Manifold Learning and Object-Oriented Data Analysis ●▲

IMS, Section on Nonparametric Statistics
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Curve and Surface Fitting: From the Viewpoint of Manifold Recovery

Edwin Chong, Colorado State University; Hari Iyer, Colorado State University; ✱Thomas Lee, Chinese University/Colorado State University, *tlee@sta.cuhk.edu.hk*; Haonan Wang, Colorado State University

Key Words: curve estimation, local linear embedding, surface fitting

Since its inception, locally linear embedding (LLE, Roweis and Saul, 2000) has become a popular method for recovering unknown low-dimensional structures hidden in high-dimensional data. Numerous successful applications to real-life problems of LLE have been reported in the literature. However, its theoretical properties are largely unknown. In this work, we define a population version of LLE and study its theoretical and statistical properties. This is particularly important for understanding the good empirical performance of LLE.

Object-Oriented Data Analysis of Tree-Structured Objects, Random Graphs, and Manifolds

✱Haonan Wang, Colorado State University, 80526, *wanghn@stat.colostate.edu*

Key Words: non-Euclidean space, Lie group, centrality, dispersion, vectorcardiography

Object oriented data, such as tree-structured data, random graphs, manifold data and curve data, are frequently collected in many scientific studies. Traditional statistical models for multivariate data are built under Euclidean space setting. However, the elements of object oriented data analysis reside in non-Euclidean spaces such as Lie groups, or more complex spaces such as spaces of tree-structured data. For example, two blood vessel systems differ in terms of topological structures and geometric properties (i.e., overall length, number of branches, and branching orientation). A mathematical framework for statistical analysis of object oriented data, including measures of centrality, variability and a notion of curves, has been carefully developed. The methodology is illustrated through applications to the analysis of vectorcardiography data and brain blood vessel data.

Simple Statistics on Interesting Spaces: Regression Analysis on Manifolds for Computational Anatomy

✱Sarang Joshi, The University of Utah, 72 S Central Campus Drive, 3750 WEB, Salt Lake City, UT 84112, *sjoshi@sci.utah.edu*; P. Thomas Fletcher, The University of Utah; Brad Davis, The University of North Carolina at Chapel Hill

Key Words: Manifold Statsitics, Regresion Analysis, Manifold Regresion, Computational Anatomy

Regression analysis is a powerful tool for the study of changes in a dependant variable as a function of an independent regressor variable. When the underlying process can be modeled by parameters in a Euclidean space, classical regression techniques are applicable and have been studied extensively. However, recent work suggests that attempts to describe anatomical shapes using flat Euclidean spaces undermines our ability to represent natural biological variability. In this talk, I will develop a method for regression analysis of general, manifold-valued data. Specifically, we extend Nadaraya-Watson kernel regression by recasting the regression problem in terms of Frechet expectation. Although this method is quite general, our driving problem is the study anatomical shape-change as a function of age from random-design image data.

212 Statistics Can Help Reduce Child Mortality ●▲

Social Statistics Section, Section on Government Statistics, Section on Quality and Productivity, Section on Survey Research Methods, Scientific and Public Affairs Advisory Committee, Section on Health Policy Statistics, Section on Statistics in Epidemiology, *CHANCE*
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Child Survival Theory and Practice: How Do They Match?

✱Gareth Jones, Adeni Consulting, 301 Parnell Ave., Ottawa, ON K1Z 5J7 Canada, *gj1@adeni.ca*

Key Words: child survival, interventions, global

The 2003 Lancet paper on child survival reviewed interventions feasible for delivery at high coverage in low-income settings. It showed that if key interventions were universally available, over 60% of then current global under-5 deaths could be prevented. Furthermore, while health facility-centric interventions are necessary, household-based interventions, with limited need for external material inputs, could prevent over half these deaths. This is the theory, what is the practice? A child survival project in four countries in West Africa serves as the basis for assessing the situation, not only in terms of intervention coverage and mortality measures, but also the more qualitative aspects of people motivation and management, logistics, adaptation to local conditions, and impact on policies. Some of the challenges of interpreting such information are considered.

Child Mortality: What We Count Counts

✱Jef L. Leroy, National Institute of Public Health, Av. Universidad No. 655, Col. Santa María Ahuacatitlán, Cuernavaca, 62508 Mexico, *jleroy@correo.insp.mx*; Jean-Pierre Habicht, Cornell University; Juan Rivera, National Institute of Public Health

Key Words: program evaluation, effectiveness, child mortality, delivery, utilization

Mortality in children (10 million annually) remains unacceptably high. Children can be saved by improving medical technology (MT), by improving delivery and utilization (DU) of the technology or by a combination of both. (1) We estimated that MT could reduce child mortality by 22%. This reduction is 1/3 of what could be achieved if DU of existing technologies were adequate. (2) We found a serious discrepancy between research needs and research funding: a mere 3% of research grants funded by NIH and the Gates Foundation were directed toward DU; 97% were related toward MT, with its relatively low potential to save lives. (3) We estimated the millions of lives currently saved by public health programs. Conclusion: effectively implemented programs could save millions of lives. Statistics plays a key role in the design and evaluation of effective programs.

How Can Statistics Save Lives?

✱Megan Price, Emory University, 1518 Clifton Rd, Biostatistics Department, Third Floor, Atlanta, GA 30322, *meprice@sph.emory.edu*

Key Words: statistics, human rights, right to health, service delivery research

Numbers can't provide food, shelter, medical treatment, or security. But well-executed statistical analyses can help to identify programs that are successful in providing these basic needs under challenging circumstances—poverty, corruption, social and political instability, etc. Although the Health and Human Rights Report Card is a work in progress, and is admittedly quite limited in scope, it does provide a starting point to identify states that appear to be out-performing their economic peers on key health indicators. Further research could then be conducted within these states to uncover successful strategies for fulfilling the right to health at high levels of population coverage under poor economic conditions. More generally, the report card should be viewed as an example of how statistics can be used to identify the most effective strategies for improving health outcomes.

Corruption Networks: Why Millions of Children Are Dying Needlessly

✱David J. Fitch, Universidad Rafael Landívar, Instituto de Investigaciones Económicas y Sociales, Guatemala, International 01901 Guatemala, *dfitch7@yahoo.com*

Key Words: Evaluation, Corruption, Developing countries, USAID, LQAS, Service delivery research

We statisticians, with our expertise in evaluation, could contribute much to the reduction of child deaths in the poor world, but there are problems. Evidence suggests that if a set of feasible, low cost health services were universally available, six million fewer children would not each year die. What is needed are statistically knowledgeable donors, prepared to stick with things over the long run, working with respected country groups, doing the needed research. The two would develop services coverage guided by experimental evaluations, similar to that carried out by PROGRESA in Mexico. But such evaluation is foreign to the donor community. I will try to spell this out, from my 18-year home in a poor indigenous village of the Guatemalan Highlands, from watching the work of USAID in the country, as a statistician in a WHO institute, and with a small, high-quality university.

213 Theory and Applications of Generalized Confidence Intervals ●

Section on Physical and Engineering Sciences, IMS
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Generalized Fiducial Inference

✱Jan Hannig, The University of North Carolina at Chapel Hill, Department of Statistics and Operations Research, Chapel Hill, NC 27599, *jan.hannig@colostate.edu*

Key Words: Fiducial inference, generalized confidence intervals, asymptotics, MCMC

In this paper we extend Fisher's fiducial argument and Weerahandi's generalized inference to obtain a generalized fiducial recipe that greatly expands the applicability of both fiducial ideas and generalized inference. We do this assuming as little structure as possible. We demonstrate generalized fiducial inference on examples and investigate, by simulation and by theoretical considerations, some properties of the statistical procedures derived by the generalized fiducial recipe.

Generalized Confidence Intervals for Normal Orthant Probabilities

✱Thomas Mathew, University of Maryland Baltimore County, Department of Mathematics and Statistics, 1000 Hilltop Circle, Baltimore, MD 21250, *mathew@math.umbc.edu*; David Webb, U.S. Army Research Laboratory

Key Words: Bivariate normal orthant probability, Generalized pivotal quantity, Tolerance limit

For the interval estimation of a bivariate normal orthant or rectangular probability, the generalized confidence interval idea will be applied. The problem has applications in ballistics, where it is required to estimate the hit probabilities based on live-fire weapon performance. This parameter is obviously a function of the unknown mean vector and covariance matrix. A second problem to be discussed is on the computation of an upper tolerance limit for the ratio of the bivariate normal random variables. This problem can be reduced to that of computing an upper confidence limit for an appropriate bivariate normal orthant probability. The problem has relevance in bioassay applications. Numerical results on the performance of the proposed generalized confidence intervals will be reported, along with some data analysis.

Using Generalized Confidence Intervals To Estimate Misclassification Rates of Bullet Lead Data

✱Connie Borror, Arizona State University West, 4506 W. Kaler Circle, Glendale, AZ 85301, *cborror@asu.edu*

Key Words: Gauge R&R Study, Measurement System, Classification

To determine the ability of a measurement system, it is often necessary to conduct a gauge repeatability and reproducibility (R&R) study. By comparing the relative magnitudes of the sources of variation due to the measurement system with the variation inherent in the measured process, decisions are made concerning the capability of the measurement system. The purpose of most measurement systems is to discriminate between good and bad items. In a more general case, this is true with any classification process. In this paper, measurement methods are proposed to classify bullet lead by their chemical composition. Using a generalized confidence interval approach previously applied to a gauge R&R applications, upper bounds on misclassification rates of these results are provided and discussed.

Assessing Equivalence of Two Assays Using Sensitivity and Specificity

✱Jorge Quiroz, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033, *jorge.quiroz@spcorp.com*

Key Words: Eqvuivalence studies, Generalized confidence intervals, Intersection-union tests, Misclassification rates, Sensitivity and specificity, Variance components

The equivalence of two assays is determined using the sensitivity and specificity relative to a gold standard. The equivalence testing criterion is based on a misclassification rate proposed by Burdick et al. (2005) and the intersection-union test (IUT) method proposed by Berger (1982). Using a variance components model and the IUT methods, we construct bounds for the sensitivity and specificity relative to the gold standard assay based on generalized confidence intervals.

214 Recent Developments for Nonpharmaceutical Clinical Trials ●▲

WNAR, Biopharmaceutical Section, Biometrics Section

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Multicenter Studies of Diagnostic Imaging

Constantine Gatsonis, Brown University; ✱Mei Hsiu Chen, Brown University, Center for Statistical Sciences, Box G-S121-7, Providence, RI 02912, *chenmei@stat.brown.edu*

Key Words: diagnostic imaging, multi-center studies, diagnostic accuracy, prediction

Although information from imaging has long been an integral part of health care, the evaluation of imaging modalities in the multicenter setting is a relatively recent development. Such studies assess the performance of imaging modalities for early detection, diagnosis, staging, and for monitoring and management of patient therapy. In this presentation we will discuss the design, conduct and analysis of multicenter studies of imaging modalities. We will highlight issues specific to imaging and will place emphasis on the most recent types of studies evaluating imaging as a biomarker for disease progression and response to therapy. The studies of imaging in cancer conducted by the American College of Radiology Imaging Network will be used as the main source of examples.

Study Designs for Biomarker-Based Treatment Selection

✱Amy Laird, University of Washington, Seattle, WA 98195, *amylaird@u.washington.edu*; Andrew Zhou, University of Washington

Key Words: Biomarker, Personalized medicine, Clinical trials, Sample size calculation

Among patients with the same clinical disease diagnosis, response to a treatment is often quite heterogeneous. For many diseases this may be due to molecular heterogeneity of the disease itself, which may be measured via a biomarker. In this talk, we consider the problem of evaluating clinical trial designs for drug response based on an assay of a predictive biomarker. In the planning stages of a clinical trial, one important consideration is the number of patients needed for the study to be able to detect some clinically meaningful difference between two treatments. We outline several marker validation trial designs in terms of the scientific questions each one is able to address, and compute the number of patients required for each one. We exhibit efficiency graphs for several special cases to summarize our results.

Subgroup Analysis for Assessing Treatment Efficacy with a Survival Outcome and Missing Subgroup Status, with Applications to Cancer Management and Clinical Trials with Noncompliance

✱Gang Li, University of California, Los Angeles, 2868 Nicada Drive, Los Angeles, CA 90077, *vli@ucla.edu*

Subgroup analysis is often needed in medical studies when only for a subgroup of the study subjects are thought to benefit from the treatment under study. For instance, when studying the efficacy of a drug in the presence of noncompliance, only the compliers who actually take the drug are expected to benefit from the drug. In this research, we develop large sample nonparametric procedures for comparing subgroup survival curves when the subgroup status is unknown in the control arm. Small sample performances of the proposed methods are investigated using simulations. The methods are illustrated using a dataset from a Multicenter Selective Lymphadenectomy Trial (MSLT).

215 New Directions in Safety Planning and Analysis for Clinical Development ●

Biopharmaceutical Section, WNAR, Biometrics Section
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Regulatory Perspectives on Planning for Pre-Marketing

✱George Rochester, U.S. Food and Drug Administration, 10903 New Hampshire Ave, Silver Spring, MD 20993, *george.rochester@fda.hhs.gov*

Key Words: risk, safety, safety analysis plan, risk identification model

In the current climate of heightened expectation from consumers, legislative bodies and other stakeholders, there is a need to improve pre-marketing risk assessment. Risk evaluation must be systematically considered throughout the lifecycle of a product. This presentation will discuss two new ideas. The first is the Pre-approval Risk Identification Model (PRIM), which aims to streamline safety evaluation throughout the life-cycle to meet regulatory requirements for good safety assessment which forms the basis for good risk management. The second is the Safety Analysis Plan, presented as a tool for documenting the risk identification processes and analytical framework for regulatory submissions, data analysis and presentation. A case study will be presented.

Planning for Meta-Analysis

✱Jesse A. Berlin, Johnson & Johnson Pharmaceutical R&D, LLC, 1125 Trenton-Harbourton Road, PO Box 200; Mail Stop 67, Titusville, NJ 08560, *jberlin@prdus.jnj.com*

Key Words: meta-analysis, safety data

Considering a clinical development program as a whole, the Integrated Summary of Safety (ISS), essentially a meta-analysis of individual patient data, should be planned to address safety questions of interest. Planning should involve consideration of statistical power and precision, anticipated subgroup analyses, and how study design elements will complement each other across the program. Planned meta-analyses offer precision for subgroup analyses that won't be available within individual studies. With regard to design elements, the issue of confounding of study design elements should be explicitly addressed. For example, a low dose study in one subgroup (e.g., women) and a high dose study in another (e.g., men), completely confounds dose and sex, making it impossible to disaggregate the effects of either on study findings.

Safety Analyses: Thoughts on Tackling Multiplicity

✱Devan V. Mehrotra, Merck Research Laboratories, Mailstop: UG1CD-44, 351 N. Sumneytown Pike, North Wales, PA 19454, *devan_mehrotra@merck.com*; Joseph Heyse, Merck Research Laboratories

Key Words: Adverse experiences, Clinical Trials, False discovery rate, Multiplicity, Safety analyses

Adverse experience (AE) data in randomized clinical trials are routinely evaluated using between-group p-values for every AE encountered within each of several body systems. If the p-values (or confidence intervals) are interpreted without multiplicity considerations, there is a potential for excessive false positive findings that can needlessly cloud the safety profile of the drug or vaccine. We propose a method for addressing multiplicity that achieves an ICH-recommended balance between type I and type II errors. The method involves a two-step application of the Benjamini and Hochberg false discovery rate (FDR) procedure. Data from three moderate to large vaccine trials are used to illustrate the proposed "Double FDR" approach, and a fourth example serves to reinforce the potential consequences of failing to account for multiplicity. (Joint work with the late Professor John W. Tukey.)

Detecting Safety Signals in Clinical Trials: A Bayesian Perspective

✱H. Amy Xia, Amgen, Inc, One Amgen Center Dr, MS 24-2-A, Thousand Oaks, CA 91320, *hxia@amgen.com*; Haijun Ma, Amgen, Inc; Bradley P. Carlin, The University of Minnesota

Key Words: multiplicity, signal detection, clinical trials, Bayesian hierarchical modeling, drug safety

Detection of safety signals from routinely collected adverse event data in clinical trials is critical in drug development. How to deal with the multiplicity issue and rare adverse event (AE) data in such a setting is a challenging statistical problem. Bayesian hierarchical mixture modeling [Berry and Berry (2004)] is appealing in the following aspects 1) it allows for explicitly modeling the AE data with the existing coding structure; 2) it is attractive in dealing with rare AE data because the model modulates the extremes; and 3) it is flexible to assess the posterior probability of a clinically important difference with different scales. In this presentation, we illustrate the use of Bayesian hierarchical binomial and Poisson mixture models for binary and subject-year adjusted outcomes, respectively. We also show some effective graphics for displaying flagged signals.

216 Nonparametric Classification

Section on Nonparametric Statistics, IMS, Section on Physical and Engineering Sciences
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Nonparametric Approach to Assessing the Performance of Classifiers

✱Peter G. Hall, The University of Melbourne, Department of Mathematics and Statistics, Melbourne, International 3010 Australia, *p.hall@ms.unimelb.edu.au*

The talk will discuss nonparametric approaches to determining the performance of statistical classifiers, in settings where dimension is high and the number of data is small.

Bayesian Ensemble Modeling for Classification Problems

✱Robert McCulloch, The University of Chicago, 5807 S. Woodlawn Avenue., Graduate School of Business, Chicago, IL 60637, *robert.mcculloch@chicagogsb.edu*

Key Words: prediction, MCMC

In "BART" (Bayesian Additive Regression Trees) Chipman, George, and McCulloch developed a fully Bayesian approach to the model: y = f(x) + e, where the errors may be drawn from any symmetric distribution. In the spirit of "ensemble models" the unknown function f was modeled as the sum of many simple tree models. The contribution of each tree was kept small through the use of a strong regularization prior. The BART methodology was shown to be very competitive in terms of out-of-sample prediction. However, the BART model, prior, and MCMC algorithm are all geared toward the case where the response is numeric. In this talk, we explore the use of the BART methodology in classification problems. Different approaches to extending BART to classification are discussed.

Data Piling Direction in High-Dimension, Low Sample Size Data

James Marron, The University of North Carolina at Chapel Hill;
✱Jeongyoun Ahn, The University of Georgia, *jyahn@uga.edu*

Key Words: Classification, High Dimension Low Sample Size, Fisher's Linear Discrimination, Maximal Data Piling, Support Vector Machine

In HDLSS classification, projecting data points onto the normal direction vector of a separating hyperplane often gives useful information about the classifier. We present an interesting discriminant direction vector called maximal data piling (MDP), onto which data points project to only two distinct values, one for each class. It only exists in general when the dimensionality of the data is larger than the sample size. There exist many such directions in the data space, however, the MDP direction has the longest distance between the projections. We investigate mathematical and geometrical properties of the MDP direction and compare its performance as a classification method with other methods such as Fisher's linear discrimination and support vector machine.

Bayesian Curve Classification Using Wavelets

✱Bani K. Mallick, Texas A&M University, Department of Statistics, TAMU3143, College station, TX 77845, *bmallick@stat.tamu.edu*; Xiaohui S. Wang, The University of Texas-Pan American; Shubhankar Ray, Merck Research Laboratories

Key Words: Functional Data, Logit Link, MCMC, wavelets, Logistic Classification

We propose classification models for binary and multicategory data where the predictor is a random function. We use Bayesian modeling with wavelet basis functions which have nice approximation properties over a large class of functional spaces and can accommodate a variety of functional forms observed in real life applications. We develop an unified hierarchical model to encompass both the adaptive wavelet based function estimation model as well as the logistic classification model. These two models are coupled together to borrow strengths from each other in this unified hierarchical framework. The use of Gibbs sampling with conjugate priors for posterior inference makes the method computationally feasible. We compare the performance of the proposed model with other classification methods such as the existing naive plug-in methods by analyzing simulated and real data sets.

217 Accurate Elections: The Role of Statisticians ●▲

Scientific and Public Affairs Advisory Committee, Section on Survey Research Methods, Social Statistics Section, *CHANCE*

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Designing an Audit System To Increase Voter Confidence in Elections

✱Michael W. Traugott, The University of Michigan, 4230 ISR, 426 Thompson Street, Ann Arbor, MI 48106-1248, *mtrau@umich.edu*

Key Words: election audit, public policy

This paper will report on preliminary results from a research project to develop empirical information to assist Michigan in the development of an election audit system. One purpose of an audit is to ensure that the electoral system is functioning in a way that provides accurate counts. Another purpose is to increase and maintain public confidence in the electoral system, suggesting that the audit should be regular and ongoing rather than episodic. The study is being conducted in three carefully arranged phases, each successive one building upon prior work and involving discussion and consultation with local election officials. This paper will report on the second phase, involving the secondary analysis of ballots from the 2006 election to simulate the results of different sampling approaches to the auditing function.

Engaging the Unengaged Voter: Vote Centers and Voter Turnout

Robert M. Stein, Rice University; ✱Greg Vonnahme, Rice University, Dept. of Political Science, MS 27, Houston, TX 77251, *vonahm@rice.edu*

Key Words: Voter turnout, Match sample

Previous election reforms designed to increase turnout have often made voting more convenient for frequent voters without significantly increasing turnout among infrequent voters. A recent innovation—Election Day vote centers—provides an alternative means of motivating electoral participation among infrequent voters. Election Day vote centers are non-precinct based locations for voting on Election Day. The sites are fewer in number than precinct-voting stations, centrally located to major population centers (rather than distributed among many residential locations), and rely on county-wide voter registration databases accessed by electronic voting machines. Voters in the voting jurisdiction (usually a county) are provided ballots appropriate to their voter registration address.

Residual Voting: A New Diagnostic

✱Paul Gronke, Reed College, 3203 SE Woodstock Blvd, Portland, OR 97202, *p_gronke@hotmail.com*

Key Words: residual voting, voting technology, elections, election reform

This paper is a preliminary report on a project developing a new diagnostic measure for residual voting (a measure that reflects the level of voting errors—both human and machine caused—that occur in an election system). Previous studies relied on county level data and did not compare residual voting across modes (election day, absentee, early). This study examines residual voting at the precinct level and across modes, allowing a far more detailed examination of the causes and consequences of residual voting. The study relies on precinct level voting data from Florida, and may include data from Ohio, subject to availability.

Measuring Voting System Failures: Survey Evidence of the Frequency of Voting Problems in the 2006 and 2008 U.S. Election

✱Stephen Ansolabehere, Massachusetts Institute of Technology, 77 Massachusetts Avenue, E53-449, Cambridge, MA 02139-4307, *sda@mit.edu*; Charles H. Stewart, III, Massachusetts Institute of Technology

We present the results of two nationwide surveys in the United States, one from the 2006 general election and one from Super Tuesday 2008. The surveys involved large national samples to allow measurement of what are low-probability events. We measure the incidence of problems relating to voter registrations, voting equipment difficulties, lines, polling place operations, and poll workers, as well as measures of the reasons for nonvoting (a variant of the Current Population Survey question). No problem affects more than a small percent of voters, but about 1 in 5 voters experience at least one problem. Among other findings: administration of voter identification differs with race of respondent; use of electronic voting equipment and requirement for voter identification leads to longer lines; and a small percent express dissatisfaction with poll workers and polling place operations.

218 Statistics-Geography Mashups on the Web ●▲

Section on Statistical Graphics
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Geo-Mashups: Using Google Earth for Data Exploration

✱Jason Dykes, City University London, Dept. Information Science, Northampton Square, London, International EC1V 0HB United Kingdom, *jad7@soi.city.ac.uk*

Key Words: Geovisualization, Geobrowser, Interactive Cartography, Exploratory Data Analysis, GIS, Mashup

Increasing volumes of geographic data and metadata give rise to opportunities to explore the geography of new data sets and the need for a common geographic framework to spatially filter data; spatially relate diverse data sets; identify spatial trends. 'Geobrowsers' provide exploratory interfaces to rich spatial data sets. Those based upon well documented data formats and with flexible applications programming interfaces offer considerable scope for exploring data through 'mashups' that combine data through such a geographic framework. Google Earth is used to develop 'geo-mashups' in which census data, mobile application use, and user content made available via the geographic Web are explored through interactive visual approaches. A number of novel techniques are presented including tag maps and tag clouds, data dials, and links between abstract and realistic graphics.

Web-Based Geovisualization and Geocollaboration

✱Alan M. MacEachren, The Pennsylvania State University, GeoVISTA Center, Dept. of Geography, 302 Walker, University Park, PA 16802, *maceachren@psu.edu*; Anthony Robinson, The Pennsylvania State University; Robert Roth, The Pennsylvania State University; Etien Koua, Disaster Preparedness and Prevention Agency; Eugene J. Lengerich, The Pennsylvania State University

Key Words: geovisualization, atlas, mapping, web-based collaboration, ESDA

In this presentation, we outline and demonstrate two projects focused on web-based geovisualization and geocollaboration. First, we describe the design, implementation, and use of a web-based model atlas supporting cancer surveillance and control. This effort extends and refines exploratory spatial data analysis methods developed originally for desktop analysis by experts to a web-based Atlas prototype supporting diverse users. Second, we introduce the Geo-EXplication Portal, a web repository and interface designed to enable geocollaboration (group work on shared problems with geographic information through geographic technologies). The G-EX Portal serves 4 tasks: to learn about new geovisual tools and analytical methods, search for multimedia research content, enable collaboration about results and tools, and support peer review of research using geovisual analytic methods.

Web Visualization of Climate Observations in a Geographical Context

✱Robert R. Gillies, Utah State University, Utah Climate Center, 4825 Old Main Hill, Logan, UT 84322-4825, *Robert.Gillies@usu.edu*

Key Words: Linked Micromaps, Climate, Climate Database, Time series, COOP observations

The presentation introduces linked micromap (LM) plots for presenting climate summaries gathered from the Cooperative Observer Program (COOP) of weather observations in the U.S. A web-based version of LM plots was modified to extract COOP data from the Utah State University's climate database and present it as parallel sequences that constitute the now familiar LM plots. The application here shows LM time series plots for the State of Utah for which the user is able to drill-down to the level of the COOP station. The LM plots of temperature and precipitation involve multivariate statistical summaries that involve, for example, the mean and median of monthly or yearly measurements. Different sorting methods reveal and accentuate relationships. Furthermore, interactive settings facilitate the rapid re-expression of choice of sorting criteria.

Map Design for Diverse Worldwide Users of the Web

✱Robert M. Edsall, Arizona State University, School of Geographical Sciences, MC 70104, Tempe, AZ 85287-0104, *robedsall@asu.edu*

Key Words: culture, user-centered design, interface design, visualization, maps, interaction

Users of computers and the Internet now vary widely in terms of language, values, traditions, and other dimensions of culture. However, much of the software that allows us to interact with computers (and to some extent each other) was designed by and for Westerners in the US or Western Europe. To what extent do these designs limit the use of, and insight from, computer systems for individuals who, because of their culture, may not interact with the computer in ways that the designers expect or intend? This paper will suggest and support the notion that culture could play a larger role in the design of graphical (geo)data representations for worldwide users. This paper develops a research framework that is rich with testable and important hypotheses concerning the influence of the cultural background of target users of graphical representations of geostatistical data for the internet.

219 Weak Instrument Robust Tests in GMM and the New Phillips Curve ●

JBES-Journal of Business and Economic Statistics
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Weak Identification: Where Do We Stand? A Case Study of the Phillips Curve

✱Frank Kleibergen, Brown University, Department of Economics, Providence, RI 02912, *Frank_Kleibergen@brown.edu*; Sophocles Mavroeidis, Brown University

We survey the recent literature on weak instrument robust tests in GMM that either test the full parameter vector or just elements thereof. We use these test procedures to re-examine the evidence on the new Phillips curve model. In contrast to earlier studies, we find that U.S. postwar data are consistent both with the view that inflation dynamics are forward-looking, and with the opposite view that they are predominantly backward-looking. Moreover, the Phillips curve has become flatter recently, and this is an important factor contributing to its weak identification.

220 Models for Binomial/Multinomial Overdispersed Data ●

Biometrics Section
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

An Improved Method for the Computation of Maximum Likelihood Estimates for Multinomial Overdispersion Models

✱Jorge G. Morel, Procter & Gamble, 6280 Center Hill Avenue, MS S1-M, Cincinnati, OH 45224, *morel.jg@pg.com*

Key Words: Mixtures of multinomial distributions, Fisher scoring algorithm, Large cluster size approximation, Forest pollen count example

A novel two-stage maximum likelihood estimation procedure due to Neerchal and Morel (2005) is presented here. Maximum likelihood estimation of the parameters of multinomial overdispersed models using the classical Fisher scoring method poses certain computational challenges. In the case of the Dirichlet-multinomial model, the challenges are overcome by noting that the Fisher Information Matrix can be computed using the beta-binomial distribution. On the other hand, in the case of the Finite Mixture of Multinomials of Neerchal and Morel (1998), an approximation theorem can be used to obtain a two-stage procedure for computing the maximum likelihood estimates. Simulation results show that the two-stage procedure is faster without loosing any accuracy. The classical example on fossil forest pollen count due to Mosimann (1962) will be revisited to illustrate the new computational techniques.

A Family of Models for Analyzing Correlated/Clustered Multinomial

✱Justin T. Newcomer, University of Maryland, Baltimore County, Department of Mathematics and Statistics, 1000 Hilltop Circle, Baltimore, MD 21250, *jnew1@math.umbc.edu*

Key Words: Multinomial overdispersion models, Generalized Linear Mixed Models, Maximum Likelihood Estimation, Generalized Estimating Equations

Two likelihood models for modeling multinomial extra variation will be introduced, namely, the Dirichlet-multinomial due to Mosimann (1962) and a Finite Mixture of Multinomials due to Morel and Nagaraj (1993), and Neerchal and Morel (1998). As in the set-up of Generalized Linear Models, the expected value of the observed vector of counts depends on a known vector of covariates and an unknown vector of parameters via "link functions." Maximum likelihood estimation will be discussed for the most common links, that is, generalized logit, cumulative logit, cumulative probit, and cumulative complementary log-log. Gains in efficiency of these likelihood models over Generalized Estimating Equations will be shown via Monte Carlo simulation studies. The proposed models will be illustrated using a classical example from Mosimann (1962) and a clinical example from Lumley (1996).

On Goodness-of-Fit Test of an Overdispersion Model

✱Santosh C. Sutradhar, Pfizer, Inc., 235 East 42nd Street, MS 685-12-20, New York, NY 10017, *Santosh.Sutradhar@pfizer.com*; Nagaraj K. Neerchal, University of Maryland, Baltimore County; Jorge G. Morel, Procter & Gamble

Key Words: Goodness-of-fit, Maximum likelihood estimation, Grouped and ungrouped likelihood, Parametric bootstrapping, Overdispersion model

Goodness-of-fit (GOF) tests available in the literature have focused on testing the null hypothesis of 'no overdispersion'. In this presentation, we will consider an omnibus GOF test (Sutradhar, Neerchal and Morel 2008) for model adequacy of a specified overdispersed distribution. This test is a direct analogue of the usual Pearson chi-squared statistic, but is also applicable when clusters are not necessarily of the same size. We present the distributional properties of the proposed test and discuss its performance characteristics. We address a key assumption that the parameter estimates are based on the cell frequencies obtained by grouping individual observation to form the GOF statistic, a fact which is not often emphasized in the literature. A simulation is carried out to investigate the performance of the GOF test. The test is illustrated using data from Haseman and Soares (1976).

Modeling Overdispersion in Multilevels Using Random Effects

✱Jeffrey R. Wilson, Arizona State University, BAC565 W.P. Carey School of Business, Tempe, AZ 85287, *jeffrey.wilson@asu.edu*

Key Words: overdispersion, non-distributional, correlated, quasilikelihood

We will consider the use of hierarchical models such as nested models with covariates measured at multiple levels as is commonly found with healthcare data, and with longitudinal studies in general but without distributional assumptions at each level, Wilson and K.J. Koehler (1991). In the hierarchical structure often the correlation introduced is ignored, causing bias in parameter estimates, under estimating of variance components, and statistical tests that are often too liberal. In essence we will introduce a hierarchical generalized linear like model structure. Building off the estimation procedures commonly used for a one-level structure in quasilikelihood models with varying links we present techniques that provide consistent and reliable estimates in hierarchical structure. A simulation study is provided to investigate the performance of these estimates at the different models.

221 New Developments in Multivariate Time Series Analysis

Business and Economics Statistics Section, IMS
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Dynamic Factor Models with Block Structure

✱Marc Hallin, Université Libre de Bruxelles, ECARES, CP 114, 44, avenue Jeanne, Bruxelles, B-1050 Belgium, *mhallin@ulb.ac.be*

Key Words: time series, panel data, dynamic factor model

Macroeconometric data often come under the form of large panels of time series, themselves decomposing into smaller but still quite large subpanels or blocks. We show how the dynamic factor analysis method proposed in Forni et al. (2000), combined with the identification method of Hallin and Livska (2007), allows for identifying and estimating shared and block-specific common factors, and for a subtle analysis of interrelations within and between the blocks. The method is illustrated with an analysis of the Industrial Production Index data for France and Germany.

Semiparametric Nonlinear Vector Autoregressive Time Series Models

✱Yehua Li, The University of Georgia, 204 Statistics Building, 101 Cedar Street, Athens, GA 30602, *yehuali@uga.edu*; Marc G. Genton, University of Geneva

Key Words: Autoregressive, Climate, Multi-index, Nonlinear, Penalized Spline, Vector time seris

We develop a new class of semiparametric nonlinear vector autoregressive time series models. It has a dimension reduction flavor in that the current vector only depends on linear projections of the past vectors. The nonlinear relationship with the past lags has an additive model structure to avoid the curse of dimensionality. All the nonparametric functions in the model are univariate and estimated via P-splines. We study estimation, hypothesis testing, asymptotics, selection of the order of the autoregression and of the smoothing parameters, and nonlinear forecasting under the proposed model. We perform simulation experiments to evaluate our model in various settings. We illustrate our methodology on a climate data set and show that our model provides more accurate yearly forecasts of the El Nino phenomenon, the unusual warming of water in the Pacific Ocean.

Arc Length Tests for Equality of Autocovariances

✱Ferebee Tunno, Clemson University, Anderson, SC 29634, *otunno@clemson.edu*

Key Words: arc length, time series, stationary, autocovariance

This talk discusses the testing of equality of autocovariances for two independent stationary time series. Several methods have been proposed to accomplish this goal, and test statistics in the time and spectral domain have been devised and studied. In this talk, we take a new approach to the problem and assess autocovariance equality by comparing the sample arc-lengths from the two series.

Time Series Analysis for PARMA Sequences of Less Than Full Rank

✱Harry Hurd, University of North Carolina, Dept of Medical Statistics, Chapel Hill, NC 27599, *hhurd@email.unc.edu*

Key Words: PARMA, Periodically Correlated, Nonstationary, Time Series

In this paper we consider several aspects of time series analysis for periodic ARMA (PARMA) sequences when the observed sequence is of deficient (innovation) rank. We will begin by discussing the rank of PC sequences and of PARMA sequences in particular, where we address conditions on parameters that produce rank deficiency. Finally, we describe some of the considerations of time series analysis of rank deficient PARMA sequences. Specifically, we address the detection of rank deficiency in the time series and the estimation of parameters by likelihood methods.

Dynamic Correlations and Stochastic Volatility

✱Denis Pelletier, North Carolina State University, Campus Box 8110, Raleigh, NC 27695-8110, *denis_pelletier@ncsu.edu*; William McCausland, University of Montreal; Shirley Miller, University of Montreal

Key Words: dynamic correlations, stochastic volatility, prior distribution

We introduce strategies for eliciting prior distributions on correlation matrices. They allow the introduction of various types of structure while permitting the analytic evaluation of the prior in terms of any hyperparameters it may depend on. We then use our results to model variance matrices of financial returns as a combination of stochastic volatility models for the standard deviations and dynamic correlations based on regime switching.

222 Recent Advances in Monte Carlo Methods ●

Section on Bayesian Statistical Science, Biopharmaceutical Section, Section on Physical and Engineering Sciences, IMS

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Particle Filters for Large-Scale Dynamic Tomography

✱Yuguo Chen, University of Illinois at Urbana-Champaign, 725 S. Wright Street, Department of Statistics, Champaign, IL 61820, *yuguo@uiuc.edu*

Key Words: Particle Filter, State Space Model, Sequential Monte Carlo, Signal Processing, Data Assimilation

The image formation of a dynamic object from tomographic measurements can be formulated as a state estimation problem. Due to the size of the state vectors, standard implementations of particle filters will fail. However, tomographic problems admit a special factorization of the posterior distribution that allows a dramatic reduction in the effective dimension of the state space. We describe particle filters which make use of this factorization for large-scale tomography problems.

Experiments with Fully Automating Markov Chain Monte Carlo

✱Murali Haran, The Pennsylvania State University, 326 Thomas Building, University Park, PA 16802, *mharan@stat.psu.edu*

Key Words: Markov chain Monte Carlo, tempering, spatial models, Gaussian fields

Markov chain Monte Carlo (MCMC) algorithms provide a very general recipe for scientists to estimate properties of complicated distributions. While their use has become commonplace, users have several difficult and time consuming decisions to make: how to construct a good algorithm, deciding whether an MCMC algorithm is producing accurate estimates, and determining an appropriate length (stopping rule) for the Markov chain. We will describe some recent approaches for automating these decisions by constructing provably fast mixing MCMC algorithms, and discuss their application in the context of spatial generalized linear models.

Gibbs Sampling for a Bayesian Hierarchical Version of the General Linear Mixed Model

✱Alicia Johnson, The University of Minnesota, 2014 DuPont Ave S Apt 7, Minneapolis, MN 55405, *ajohnson@stat.umn.edu*; Galin Jones, The University of Minnesota

Key Words: convergence rates, Markov chain Monte Carlo, block Gibbs sampling, drift condition, minorization condition

We consider block Gibbs sampling for a Bayesian hierarchical version of the normal theory general linear mixed model. We establish drift and minorization conditions for this block Gibbs sampler. This implies that the underlying Markov chain converges to the posterior distribution at a geometric rate which is important for at least two reasons: (1) a Markov chain that converges quickly is crucial to achieving effective simulation results in finite time and (2) it is a key sufficient condition for the existence of a central limit theorem and for consistent estimation of Monte Carlo standard errors. Drift and minorization can also be used in conjunction with established methods to find good starting values for the simulation. Thus, practitioners using our

Gibbs sampler will have the tools to be as confident with the results as if they were able to draw directly from the posterior.

Markov Chain Monte Carlo: Can We Trust the Third Significant Figure?

✱Galin Jones, The University of Minnesota, School of Statistics, 224 Church Street SE, Minneapolis, MN 55455, *galin@stat.umn.edu*; Murali Haran, The Pennsylvania State University; James Flegal, The University of Minnesota

Key Words: Convergence Diagnostic, Markov Chain, Monte Carlo, Standard Errors

Current reporting of results based on Markov chain Monte Carlo computations could be improved. In particular, a measure of the accuracy of the resulting estimates is rarely reported. Thus we have little ability to objectively assess the quality of the reported estimates. We address this issue in that we discuss why Monte Carlo standard errors are important, how they can be easily calculated in Markov chain Monte Carlo and how they can be used to decide when to stop the simulation. We compare their use to a popular alternative in the context of two examples.

Assessing Convergence and Mixing of Markov Chain via Stratification

✱Rajib Paul, The Ohio State University, 1958 Neil Avenue, Columbus, OH 43210, *rajib@stat.osu.edu*; Steven N. MacEachern, The Ohio State University; Mark Berliner, The Ohio State University

Key Words: ergodicity, batch-mean method, delta method, functional CLT

Some posterior distributions leading to MCMC chains are naturally thought of as a collection of sub-chains. Examples include mixture models, regime-switching models, hidden Markov models, etc. We apply the notion of post-stratification to develop MCMC-based estimates. The overall estimates are obtained by combining different sub-group (sub-chain) estimates. Variance estimates of the limiting distributions of such estimators are developed. Based on these variance estimates we propose a test statistic which helps in checking convergence and mixing of MCMC. This statistic is compared to the Gelman-Rubin statistic. Our examples include a Bayesian change-point model for ice flow velocity in East Antarctica, a logistic regression model, and a latent variable model for Arsenic concentration in public water systems in Arizona.

223 Winners of the 2008 Statistical Computing and Statistical Graphics Student Paper Competition ▲

Section on Statistical Computing, Section on Statistical Graphics

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Use of Sparse Linear Discriminant Analysis in Classification and in Testing Gene Pathways

✱Michael C. Wu, Harvard University, 655 Huntington Ave., Building 2, 4th Floor, Boston, MA 02115, *mwu@hsph.harvard.edu*; Lingsong Zhang, Harvard University; Xihong Lin, Harvard University

Key Words: Linear discriminant analysis, Variable selection, Classification, Gene pathways, Proteomics

Linear discriminant analysis (LDA) has been effectively applied in many settings such as proteomics/genomics. We consider two-class sparse LDA (sLDA) which imposes an L1-constraint on the discriminant direction to incorporate variable selection and remove the contribution of noise variables. Our formulation leads to a piecewise linear solution path and we exploit this property to develop a fast computational algorithm for computing the entire regularized solution path. We study two applications. First, sLDA has high classification accuracy when predicting mercury exposure from proteomic profiles. Second, sLDA leads to a novel approach for testing differential gene pathway activity when many genes are contributing only noise and is applied to a microarray experiment studying metal fume exposure. The environmental health implications of our analyses are interesting and will be discussed.

Multi-Objective Optimal Experimental Designs for Event-Related fMRI Studies

✱Ming-Hung Kao, The University of Georgia, Department of Statistics, *jasonkao@uga.edu*; Abhyuday Mandal, The University of Georgia; Nicole Lazar, The University of Georgia; John Stufken, The University of Georgia

Key Words: compound design criterion, design efficiency, genetic algorithms

The nature of fMRI studies impels the need for multi-objective designs that simultaneously accomplish statistical goals, circumvent psychological confounds, and fulfill customized requirements. Taking advantage of well-known fMRI designs, we propose an efficient logarithmm to search for optimal multi-objective designs. This algorithm significantly outperforms previous search algorithms in terms of achieved efficiency, computation time and convergence rate.

Tableplot: A New Display for Factor Analysis

✱Ernest Kwan, York University, 4700 Keele Street, Toronto, ON M1V 4J1 Canada, *ernest@yorku.ca*

Key Words: factor analysis, tabular data, graphical display, statistical graphics

Factor analysis (FA) is an important methodology of the social sciences. Reviews of FA applications, however, reveal general misuses by researchers. These misuses can be partly attributed to the inability to interpret and convey the large quantity of parameter estimates in typical factor analyses. I propose the tableplot as a more effective means of presenting FA results. I illustrate how this new display can improve applications of FA.

Sparse Permutation Invariant Covariance Estimation

✱Adam Rothman, The University of Michigan, *ajrothma@umich.edu*; Peter Bickel, University of California, Berkeley; Elizaveta Levina, The University of Michigan; Ji Zhu, The University of Michigan

Key Words: Covariance matrix, High dimension low sample size, large p small n, Lasso, Sparsity, Cholesky decomposition

The paper proposes a method for constructing a sparse estimator for the inverse covariance (concentration) matrix in high-dimensional settings. The estimator uses a penalized normal likelihood approach and forces sparsity by using a lasso-type penalty. We establish a rate of convergence in the Frobenius norm as both data dimension p and sample size n are allowed to grow, and show that the rate depends explicitly on how sparse the true concentration matrix is. We also show that a correlation-based version of the method exhibits better rates in the operator norm. The estimator is required to be positive definite, but we avoid having to use semi-definite programming by re-parameterizing the objective function in terms of the Cholesky factor of the concentration matrix, and derive an iterative optimization algorithm that reduces to solving a linear system at each iteration.

224 Statistical Consulting in Civil Litigation ●▲

Section on Statistical Consulting, Social Statistics Section
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Measures of Discrimination and Protocols for Evidence: The Case of the 'Spare List'

✱Steven M. Crunk, San Jose State University, Mathematics Department, San Jose, CA 95192-0103, *crunk@math.sjsu.edu*; William B. Fairley, Analysis & Inference, Inc.

Key Words: discrimination, measure, legal evidence, protocol, selection, expert

The selection of measures of discrimination is a critical step in determining if there is legal evidence to establish discrimination. In a legal case that went to trial, an individual alleged his minority status was a factor in placing him on his employers "Spare List." Each expert testifying on the statistics used different measures. These different measures resulted in sharp differences in the characterization of discrimination. Grounds for selection of one measure over another are treated in connection with the case. These grounds include the stage of the case in legal proceedings (e.g., prima facie versus "business necessity" phases), the kind of discrimination alleged, and the interest attaching to frequency vis a vis the severity. We make suggestions for protocols that might be formally or informally adopted by courts or experts in choosing measures and avoiding selection biases.

Issues in the Use of Survival Analysis To Estimate Damages in Equal Employment Cases

✱Qing Pan, The George Washington University, Statistics Department, Washington, DC 20052, *qpan@gwu.edu*; Joseph Gastwirth, The George Washington University

Key Words: Damage Compensation, Discrimination, Equal employment cases, Parametric models, Predicted job tenure, Survival analysis

When plaintiffs prevail in civil cases, courts award damages to compensate their loss. In equal employment cases concerning hiring or promotion discrimination, a successful plaintiff deserves the earnings they would have received but for the discrimination. Survival analysis can be used to estimate the expected length of time a plaintiff would have remained an employee. This figure can be multiplied by the salary to obtain the lost earnings. One cannot simply run a standard survival regression on all the data because the sample of employees is a selected or "biased" sample since a court has found discriminatory practices were carried out. Following the observations made by judges in three legal cases, we model the biased process and question the way the damages were estimated in a different law case. Recommendations are made for the proper use of survival methods in this application.

Use of External Data Sources To Demonstrate Control for Potentially Confounding Factors

Rose M. Ray, Exponent-Failure Analysis Associates; ✱Madhu Iyer, Exponent-Failure Analysis Associates, 149 Commonwealth Drive, Menlo Park, CA 94025, *miyer@exponent.com*

Key Words: crashworthiness, PAR, risk, benefit, confounder, Delta-V

Traffic accident reports (PARs) provide a rich source for evaluation of the field performance of passenger vehicles. This data is used in litigation to evaluate the risks and benefits of vehicle designs. One use of this data in litigation is evaluation of crashworthiness (risk of injury to occupants in a crash). The change in velocity at impact, or Delta-V, of the vehicle in the crash is one of the most important factors influencing the risk of injury. A criticism of comparative crashworthiness studies based upon PARs is that these reports do not record the potentially confounding factor, Delta-V. While it is clear that Delta-V is related to the risk of injury, it is a confounding factor only if the distribution of Delta-V is sufficiently different among vehicle groups to be a confounder. Use of other data sources to test for confounding and the acceptance of this by courts is discussed.

Statistical Issues in Intellectual Property and False Advertising Litigation

✱Colleen Kelly, Exponent, Inc., 16450 Via Esprillo, San Diego, CA 92127, *ckelly@exponent.com*

Key Words: Civil litigation, intellectual property, trademark infringement, false advertising, product comparison

Advertising claims that one product is equivalent to or better than another name brand product may incite false advertising or trademark infringement litigation. Statistical expertise in designing and analyzing experiments to compare products is useful in this arena. We discuss statistical issues in a couple of case studies.

How Can Statisticians Help Improve the Quality of Construction Defect Litigation?

✱Louis A. Cox, Cox Associates, 503 Franklin Street, Denver, CO 80218, *tcoxdenver@aol.com*

Key Words: Civil litigation, construction defects, acceptance sampling, game theory, remediation, optimal statistical decisions

Construction defect litigation is increasing for homes in developments. Yet, statistical methods used to estimate damages and costs are often primitive, with plaintiffs sometimes taking ad hoc samples and defendants insisting on sample survey research designs. How can statistics help both sides? This paper views litigants and the court as playing a game where each litigant attempts to minimize its own net costs and the court seeks to minimize total social costs from unjustified remediations and failures to remediate. Efficiency can be improved for all by use of: (a) Lot acceptance sampling and adaptive sampling ideas; (b) Failure time and defect data analysis models and methods; (c) Conditional probability relations between results of visual inspections and destructive tests; and possibly (d) A priori agreement on how statistical evidence will be used in settlement negotiations.

225 Statistical Issues in Biomarker and Imaging Studies ●

Section on Statistics in Epidemiology, Biopharmaceutical Section, Biometrics Section
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Regulatory Issues Related to the Evaluation of Biomarkers in a Clinical Study

✱Kyunghee K. Song, U.S. Food and Drug Administration, 1350 Piccard Dr., CDRH, rockville, MD 20850, *kyunghee.song@fda.hhs.gov*

Key Words: biomarker, diagnostic test, clinical study, regulatory issue

There has been a great deal of interest in the use of biomarkers to drive clinical decision making for patients with cancer. Targeted therapies are often tied to a particular biomarker but how that marker is assessed using a diagnostic device can vary considerably. Most biomarkers have analytical error associated with the assay. Furthermore, by measuring DNA, gene expression (RNA) or proteins (e.g. via IHC) we are not capturing the target in a comparable manner. It is unclear how one then can tie biomarker performance to what counts, what it predicts in terms of survival endpoints for a patient. I will consider different levels of diagnostic performance and examine how this would impact the statistical evaluation of clinical outcome measures. Finally, I would like to address some regulatory concerns.

Capturing Performance of Prognostic Markers: An Overview

✱Estelle Russek-Cohen, U.S. Food and Drug Administration, 1350 Piccard Drive, HFZ-550, DBS/OSB/CDRH/FDA, Rockville, MD 20850, *Estelle.Russek-Cohen@fda.hhs.gov*; Harry F. Bushar, U.S. Food and Drug Administration; Iram Quraishi, U.S. Food and Drug Administration

Key Words: Prognostic Markers, time to event data, time dependent ROC

There have been thousands of papers identifying biomarkers that have promise as prognostic markers in cancer yet few have been cleared or approved as such by the FDA. We will present an overview of the challenge of designing studies to support a prognostic intended use. We will also describe several statistical methods that have been proposed and contrast them in relation to study design. Early stage (pre-metastatic) disease prognostic markers and late stage markers can be quite different because the utility of each may be quite different and capturing patients within a single standard therapy is almost impossible in late stage cancer.

Diagnostic Assays for Screening or Prognosis with Small Biomarker Panels

✱Friedemann Krause, Roche Diagnostics GmbH, Biostatistics, Nonnenwald 2, Penzberg, International D-82377 Germany, *friedemann.krause@roche.com*

Key Words: biomarker panels, diagnostic assays, classification, sensitivity, specificity

For detection of early diseases, single biomarkers are often lacking sensitivity and/or specificity to discriminate between diseased and non-diseased populations. Therefore, diagnostic industry pays more attention to the search and development of biomarker panels instead of a single marker. Challenges are (amongst others): to establish strategies for biomarker search, to provide qualified classification and feature selection methods for picking up the best biomarker combination, and to validate biomarker panels. By discussion of concrete examples, we show for small panels of proteomics markers possible solutions for the different steps of the process of marker panel selection and evaluation.

Existing Problems in Gene Expression Microarray-Based Assays for Clinical Practice

✱Samir Lababidi, U.S. Food and Drug Administration, 1350 Piccard Drive, HFZ-550, CDRH, Rockville, MD 20850, *samir.lababidi@fda.hhs.gov*

Key Words: Genomic classifier, gene expression, microarray, biomarker, reproducibility

Genomic classifiers using DNA microarrays are becoming powerful tools in the medical community with the potential to revolutionize the diagnosis and treatment of disease. However, despite the tremendous interest in using these classifiers in diagnosis and the management of disease, few genomic classifiers have made it into clinical practice. At the same time, the explosion of research in genomic technology is setting the ground work for these genomic classifiers to become the major tools for evaluating disease susceptibility, diagnosis, and prognosis. In this presentation we discuss specific existing problems when developing gene expression microarray-based assays during the process of validating the genomic classifier, establishing the reproducibility of the assay, and showing the added benefit of microarray experiment over clinical covariates alone for their use in clinical practice.

Maximum Likelihood Estimators for Fraction of False Negative Lesions when Radiologist Panel Is Non-Unanimous

✱Arkendra De, U.S. Food and Drug Administration, *arkendra.de@fda.hhs.gov*; Gene Pennello, U.S. Food and Drug Administration

Key Words: CAD, radiology, imaging

In settings where CAD (computer aided detection) devices for images are studied, the fraction of marks made by CAD which are true lesions yet missed by the radiologist reader (i.e. false negative fraction) is one metric which can be used to quantify CAD performance. In order to determine how radiologist readers will interpret the CAD marks, a panel of radiologists (e.g., three) is used in CAD studies. Situations in which the panel unanimously agrees with the interpretation of a CAD mark can easily be utilized in the estimation of false negative fraction. We present maximum likelihood estimators for false negative fraction when a variable number of radiologists agree.

226 Statistics and Methodological Research in the Medical Expenditure Panel Survey (MEPS)

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Evaluation of Alternative Prediction Models To Oversample Low-Income Persons in the Medical Expenditure Panel Survey (MEPS)

✱Lap-Ming Wun, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *Lap-Ming.Wun@ahrq.gov*; Trena M. Ezzati-Rice, Agency for Healthcare Research and Quality

Key Words: Prediction model, complex survey

The MEPS is designed to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. Each year a new panel of households is selected from households that participated in the previous year's National Health Interview Survey (NHIS). For most years, the MEPS sample includes an oversample of targeted policy relevant subpopulations which include selected minorities and households with low income. The status of individuals' income in the next year is not known at the time the sample is drawn. However, a wide range of potential predictors are available from the NHIS to inform the oversampling of households predicted to be poor. This paper reports on the modeling procedure and the results of updating the prediction model using data from the 2003 NHIS and from the 2004 MEPS.

Comparison of Imputation Adjustment Techniques on Variance Estimation in the Medical Expenditure Panel Survey (MEPS)

Robert M. Baskin, Agency for Healthcare Research and Quality; ✱Marc W. Zodet, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *marc.zodet@ahrq.hhs.gov*; Trena M. Ezzati-Rice, Agency for Healthcare Research and Quality

Key Words: multiple imputation, replicated imputation, Rao-Shao method, complex survey data

MEPS is a national probability sample survey designed to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. Depending on the type of medical event, there are varying levels of item nonresponse on medical expenses as collected in the MEPS household interview. MEPS also collects expenditure data in the Medical Provider Component (MPC) of the survey. Missing expenditure data for health care events are completed through a weighted sequential hot deck procedure with MPC data as the primary donor source. Studies in 2004 and 2005 examined the impact of imputation on estimates of variance for MEPS health care expenditures. This study updates this research by investigating multiple imputation as a method to assess the impact of imputation on the variance estimates.

Event Reporting in the Medical Expenditure Panel Survey (MEPS)

✱Frederick C. Rohde, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *frederick.rohde@ahrq.hhs.gov*

Key Words: panel survey, simulation

MEPS is a nationally representative survey of U.S. civilian households. Respondents are asked about their medical insurance, healthcare utilization and expenses through five rounds of interviews conducted over a two-year period. Annual estimates for the US are made by combining the data from the panel in its first year with the one in its second. This research examines if the level of reporting of medical events differs between each panel's first and second years and if the drop occurs primarily between the first and second rounds compared to later rounds. An assessment of the impact of any differential reporting on the annual MEPS expenditure estimates is made using simulation.

Sample Redesign and Conversion to a Windows-Based Survey Instrument: Impact on Part-Year Estimates from the 2007 Medical Expenditure Panel Survey (MEPS)

✱Trena M. Ezzati-Rice, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *trena.ezzati-rice@ahrq.gov*; Frederick C. Rohde, Agency for Healthcare Research and Quality; Steven R. Machlin, Agency for Healthcare Research and Quality

Key Words: Longitudinal survey, panel effects, complex survey

The MEPS is a longitudinal panel survey. A new panel is fielded annually and data are collected via five CAPI interviews that cumulatively cover two consecutive years. Each annual MEPS panel is selected as a subsample of respondents to the prior year's National Health Interview Survey (NHIS). In 2006, a new sample design was implemented for the NHIS and the changes carried over to the MEPS 2007 panel. Also, in 2007 the MEPS survey instrument was converted from a DOS-based system to a Windows-based system. These two changes have the potential to affect point estimates and standard errors of MEPS estimates. In this paper, we conduct an assessment of the impact of the changes on preliminary 2007 estimates. In particular, cross-panel comparisons of health status and health insurance coverage are made for estimates from the first part of 2007.

Comparison of Direct, Mixed Model, and Bayesian Metropolitan Statistical Area Estimates for the Insurance Component of the Medical Expenditure Panel Survey

✱Robert M. Baskin, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, *robert.baskin@ahrq.hhs.gov*; John P. Sommers, Agency for Healthcare Research and Quality

Key Words: complex sample, design bias, mean squared error

The Medical Expenditure Panel Survey (MEPS) is a national probability sample survey. The Insurance Component of MEPS provides national and state level estimates of insurance coverage for the U.S. civilian noninstitutionalized population. The demand for reliable data at the state level and below, regarding healthcare insurance has greatly increased. Previous research has been conducted to produce direct design-based estimates using the MEPS Insurance Component design structure. However, the number of Metropolitan Statistical Areas (MSAs) for which direct estimates can be produced with acceptable reliability is limited. In this paper, we evaluate mixed models and Bayesian models, incorporating a time covariate, to produce MSA level estimates for smaller MSAs in ten states. We examine estimates of MSE and RSE of two types of estimates based on direct and indirect estimation techniques.

227 Spanning the Globe: The Introductory Statistics Course as a Gateway to the Rest of the World ▲

Section on Statistical Education

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Spanning the Globe: The Intro Stats Course as a Gateway to the Rest of the World

✱Robert H. Carver, Stonehill College, Dept. of Business Administration, 320 Washington Street, Easton, MA 02357, *rcarver@stonehill.edu*; ✱Nilupa S. Gunaratna, International Nutrition Foundation, 150 Harrison Avenue, Room 232, Boston, MA 02111, *ngunaratna@inffoundation.org*; ✱James J. Cochran, Louisiana Tech University, College of Business, PO Box 10318, Ruston, LA 71272, *jcochran@cab.latech.edu*; ✱Paul J. Fields, Brigham Young University, 223A TMCB, Provo, UT 84602, *pjfields@byu.edu*

Key Words: Introductory statistics, Statisticst Education, Globalization, GAISE, Real data, Study abroad

As the world becomes "flatter" (in Thomas L. Friedman's phrase), the undergraduate introductory statistics course is in a unique position to expand the world view of our students by connecting them to issues and conditions around the world. Panelists will review options for accessing international sources of data, share experiences of a study abroad program for statistics students, consider the cultural challenges of teaching statistics internationally, and discuss ways of framing substantive course examples around global demography, public health, development, education, economics, and social justice.

228 Ideas for Improving Statistical Competence of Health Science Professionals and Graduate Students

Section on Teaching Statistics in the Health Sciences, WNAR, Section on Statistics in Epidemiology
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Statistical Competence or Statistical Literacy: What Should We Teach Health Professionals?

✱Janet Tooze, Wake Forest University Health Sciences, Medical Center Blvd., Winston-Salem, NC 27106, *jtooze@wfubmc.edu*

Key Words: Education, Teaching, Health Sciences

Many statistics courses for health professionals focus on teaching statistical competence (i.e., teaching how to conduct basic statistical tests, such as contingency table methods, t-tests, regression, and ANOVA). In the liberal arts, there has a call for teaching students statistical literacy (i.e., teaching general statistical thinking and reasoning as opposed to how to use specific data analytic tools). This paper will discuss the similarities and differences of the needs of health professionals (primarily MDs) compared to students receiving a liberal arts education, and frame the concepts of statistical competence and statistical literacy for this audience. Results from a survey of the statistical competence and statistical literacy of current and former masters' students in a clinical research training program will be presented.

The Scientist Game: Power and Subterfuge in the Statistical Design of Studies

✱Scott S. Emerson, University of Washington, Box 357232, Seattle, WA 98195, *semerson@u.washington.edu*

Key Words: Clinical trial design, statistical power

Proper characterization of the hypotheses addressed by a randomized clinical trial or other biomedical study is crucial for evaluation of the study's scientific merit and for ensuring the ethical treatment of study subjects. The most standard approaches presented in introductory statistics texts often do not adequately address the true scientific issues. Drawing on the game of Eleusis as described by Martin Gardner in Scientific American 30 years ago, I developed an exercise ("The Scientist Game") to illustrate some common foibles of scientists and statisticians when they first approach the statistical design of an experiment. This example is used to illustrate the importance of clearly defining relevant and important hypotheses, mapping out a scientific strategy that extends beyond a single experiment, and properly powering the study or studies to discriminate between hypotheses.

Quantitative Literacy Among Health-Care Professionals

✱Mary Z. Mays, College of Nursing & Healthcare Innovation, Arizona State University, 500 N 3rd Str, Phoenix, AZ 85004, *mary.mays@asu.edu*; Jan Jirsak, College of Public Health, University of Arizona

Key Words: statistical education, quantitative literacy, healthcare, prevalence, NAAL, numeracy

Background: Extensive research exists on the health disparities associated with limited literacy among patients. The prevalence of limited literacy in patients is 30% to 50%. Objective: Examine whether healthcare professionals have sufficient quantitative literacy (QL) to assist patients with limited literacy. Method: Data from the National Assessment of Adult Literacy (N > 19,000; weighted N = 212,819,074) were cross-tabulated by educational attainment and occupation. Results: The prevalence of "proficient" QL (score >= 350/500, 70% correct) among professionals with post-baccalaureate education was 34% (SE = 12) in health diagnosis/treatment occupations, 21% (23) in healthcare technology/support occupations, and 37% (3) in all other occupations combined. Conclusion: Graduate education in healthcare may not provide clinicians with the QL needed to assist patients with limited literacy.

Team-Based Learning in a Graduate Statistical Methods Course for Biomedical Researchers

✱Jessica M. Ketchum, Virginia Commonwealth University, Biostatistics 980032, R, 23298-0032, *MckinneyJL@vcu.edu*; Al M. Best, Virginia Commonwealth University; Emily H. Sheldon, Virginia Commonwealth University

Key Words: team-based learning, active-learning, teaching, statistical methods, graduate

How material is presented can have a dramatic impact on student learning. Active learning strategies, such as team-based learning, have been shown to be more effective than more traditional lectures. In a graduate Biostatistics course four TBL activities were implemented for two primary types of exercises: 1) perusal of journal articles and 2) review of material presumed to be prerequisite for the course. Students were given a reading exercise prior to class. During class, students individually took a short quiz and then worked within teams to repeat the quiz. Finally, the class discussed any remaining difficult or confusing concepts. A weighted average of the quizzes accounted for 5% of the course grade. The amount and quality of the discussion within class improved as compared to prior years. Furthermore students preformed significantly better within teams as compared to individually.

Teaching Bayesian Statistics in a Health Research Methodology Program

✱Eleanor M. Pullenayegum, St. Joseph's Healthcare, FSORC, 50 Charlton Ave. E, Hamilton, ON L8N 4A6 Canada, *pullena@mcmaster.ca*; Lehana Thabane, McMaster University

Key Words: Biostatistics, teaching, inter-disciplinary, Biostatistics

Although the biostatistics literature is rich in Bayesian methods for clinical trials, the use of these methods has been limited, particularly amongst academic clinicians. We believe that this is in part due to a lack of effective communication between biostatisticians and epidemiologists. Moreover, while there are courses that train biostatisticians in the theory of Bayesian methods there are relatively few that effectively train biostatisticians in the practical use of Bayesian methods. Similarly graduate courses that train epidemiologists in Bayesian methods are rare. We describe our experience of developing and teaching a course in practical Bayesian design and analysis of clinical studies, intended for a mixed group of biostatisticians, epidemiologists and clinician researchers enrolled in a Health Research Methodology graduate program.

Using the Motivation of Graduate Students To Write Defensible Theses To Teach Biostatistical Methods

✱Al M. Best, Virginia Commonwealth University, Biostatistics 980032, Richmond, VA 23298-0032, *ALBest@VCU.edu*; Jessica M. Ketchum, Virginia Commonwealth University; Emily H. Sheldon, Virginia Commonwealth University

Key Words: Statistical methods, Teaching, Online course materials, Writing intensive

Most biomedical graduate programs require students to defend their statistical methods when presenting their thesis. We prepare them for this by using real data, the GAISE principles, and intensive writing assignments. They choose an appropriate statistical method for the question at hand; produce statistical analyses to answer the question; and they document their understanding using the style of a results section of a journal article. Assessments are designed to give feedback on how to improve their writing. Lang & Secic's "How to Report Statistics in Medicine" is used to describe the style of exposition used in scientific writing. Increased course enrollment and high course evaluations show positive results. Students increase their confidence in their ability to step thru the process of turning raw information into a defensible narrative that expresses their understanding.

229 Bootstrap and BRR Variance Estimation

Section on Survey Research Methods, Section on Government Statistics, Section on Statistical Computing, SSC

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Variance Estimation for Statistics Canada's Small Household Surveys in the Context of the Household Survey Strategy

✱Sebastien Landry, Statistics Canada, 100 Tunney's Pasture Driveway, R.H.Coats Building, Floor 16-R, Ottawa, ON K1A 0T6 Canada, *sebastien.landry@statcan.ca*

Key Words: Variance estimation, Multi-phase surveys, Multi-frame surveys, Bootstrap, Household surveys

Statistics Canada launched the Household Survey Strategy (HSS) to be able to increase its capacity to conduct ongoing and new household surveys in a cost-effective manner. The main methodological component of the HSS is the master sample, which will group the respondents from major household surveys to produce a sample frame which will be used by smaller household surveys to select their samples. Many methodological challenges will confront these multiphase, multistage and multiframe smaller surveys, including sample coordination, weighting and estimation. This paper will focus on the variance estimation for the smaller surveys. A new variant of the bootstrap method is planned to be used to estimate the variance of these surveys. A simulation study trying to recreate the HSS environment using census data shows that this method should give fairly accurate variance estimates.

Implementing Resampling Methods for Design-Based Variance Estimation in Multilevel Models: Using HLM6 and SAS Together

✱Fritz Pierre, Statistics Canada, 100 Tunney's Pasture DW, Ottawa, ON K1A 0T6 Canada, *fritz.pierre@statcan.ca*; Abdelnasser Saïdi, Statistics Canada

Key Words: Multilevel modeling, Resampling methods, Design-based variance, HLM6

Multilevel models (MLM) are often used to analyze clustered data. Most software packages developed for fitting such models have implemented and recommend the use of a robust sandwich estimator to estimate the design-based variance (DBV) of the model parameters. However, this method is known to underestimate the DBV due to the plug-in principle and by possibly ignoring the stochastic adjustments to the sampling weights (Kovacevic et al. (2006)). Therefore, resampling methods have been proposed to more accurately estimate the DBV. However, no software that can currently fit MLM to survey data has implemented these techniques. In this paper, we present a review of the DBV estimation methods for estimated parameters in MLM. We then describe BHLMSAS_V0, a SAS macro that has been developed to use HLM6 (specialized software to fit MLM) in the estimation of the DBV by resampling methods.

Performance of Bootstrap Variance Estimation for a Dual Frame Household Survey: Evidence from the German Panel Survey

✱Hans Kiesl, Institute for Employment Research, Regensburger Str. 104, Nuremberg, International 90478 Germany, *hans.kiesl@iab.de*

Key Words: variance estimation, dual frame survey, bootstrap

The German Institute for Employment Research has just completed the first wave of a new panel survey focusing on low income households. It is designed as a dual frame survey, the first frame being the register of households receiving some kind of unemployment benefits, the second frame consisting of an address register of the whole population. 6000 households were selected from each frame (with stratified PPS sampling of zip codes in the first stage), resulting in a large variance of design weights between the two samples. In this paper, we describe the sampling design and the construction of final (single frame) weights, and we show the results of a simulation study that tries to evaluate the performance of bootstrap variance estimators for our survey.

Bias of BRR Variance Estimation in Surveys Weight Adjusted for Nonresponse

✱Eric V. Slud, U.S. Census Bureau, Statistical Research Division, 4600 Silver Hill Rd., Rm. 5K004, Washington, DC 20233-9100, *Eric.V.Slud@census.gov*; Yves Thibaudeau, U.S. Census Bureau

Key Words: Variance estimation, Replication methods, Weight Adjustment, Misspecified model, Quasi-randomization, SIPP

Nonresponse adjustments in large complex surveys are typically done by adjusting weights in one or more steps of ratio-adjustment to reflect observed rates of response within geographically and demographically defined population cells. Variances of survey estimates are then often estimated by replication methods (BRR). Under a quasi-randomization model for nonresponse, we extend variance formulas of Sarndal and Lundstrom (2005) to cases where the cell adjustment models are misspecified and compare them with the large-sample limits of BRR-type variance estimators. Discrepancies between the true and BRR-estimated variances are seen to arise from half-PSU imbalances, from systematic PSU differences in true versus adjustment-cell modeled response probabilities, or from PSU differences in attribute means. Results are illustrated with real and simulated data from SIPP.

Application of Fay's Method for Variance Estimation in the National Compensation Survey Benefits Products

✱Jonathan J. Lisic, Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 3160, Washington, DC 20212, *lisic.jonathan@bls.gov*; Lola Ojo, Bureau of Labor Statistics

Key Words: replication, multi-stage sampling, variances

Data on employee benefits is one of the key products produced by the National Compensation Survey (NCS). The NCS uses a stratified three-stage sample: areas (PSUs), establishments, and occupations. The survey weighting procedure includes adjusting sample weights for establishment and occupation nonresponse, unusual data collection situations, and post-stratifying to known employment control totals. This paper illustrates how Fay's method of balanced repeated replication (BRR) was applied to estimates generated from this complex sample design to obtain variance estimates for the NCS benefits products. The following topics will be presented: an introduction to the NCS survey; a description of our utilization of Fay's method, including formation of variance strata, areas (PSUs), and replicates; a detailed analysis of the magnitude of estimated benefits variances.

Using Post-Stratification To Adjust Horvitz-Thompson Estimation and Balanced Repeated Replication for Nonresponse in Longitudinal Surveys

✱Yves Thibaudeau, U.S. Census Bureau, 1037 17th St S, Arllington, VA 22202, *yves.thibaudeau@census.gov*; Eric V. Slud, U.S. Census Bureau

Key Words: BRR, Method of Laplace, SIPP, Nonresponse Adjustment

The Survey of Income and Program Participation, conducted by the U.S. Census Bureau, is a longitudinal survey plagued by severe cases of missing data. So, the cross-sectional Horvitz-Thompson estimator for subdomain totals cannot be computed, much less its variance. One solution to this problem is to prorate partial subdomain estimators from previous cross-sections, in lieu of the missing data. The result is a "backdated" HT subdomain estimator. We use MLE's of transition probabilities (Bishop Fienberg Holland 1975 page 267) to prorate. We compute the MLE's through poststratification (Little 2004, 1993). The variance formula combines large sample approximations and BRR. This approach takes advantage of the survey design and the power of the poststratification. It avoids having to modify the replication scheme to account for nonresponse.

Bootstrap Variance Estimation for Predicted Individual and Aggregate Risks

✱Milorad Kovacevic, Statistics Canada, 15 R.H.C. Coats Building, 100 Tunney's Pasture, Ottawa, ON K1A0T6 Canada, *kovacmil@yahoo.com*; Georgia Roberts, Statistics Canada; Lenka Mach, Statistics Canada

Key Words: complex sample design, logistic regression, cross-validation

Often there is a need to predict the probability that an individual with specific characteristics (risk factors) will suffer from a certain disease or a health condition. Sometimes the real interest is to provide the aggregate predicted risk estimates at the population level or at the level of a subpopulation (e.g., males age 65+). A related problem is prediction of such a probability for an individual who does not belong to the surveyed population. While the estimation of these probabilities follows from fitting an appropriate model to the available data, the design-based standard error estimation of so obtained estimates is not obvious. We are proposing a bootstrap method for estimation of the standard errors of predicted individual and aggregate risks. The method is illustrated using data from the Canadian Community Health Survey.

230 Rankings, Groupings, and Predictions in Sports

Section on Statistics in Sports

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

The 'Pablo' Volleyball Ranking System

✱Paul G. Wenthold, Purdue University, Department of Chemistry, 560 Oval Drive, West Lafayette, IN 47907, *pgw@purdue.edu*

Key Words: rankings, volleyball

This paper describes a volleyball ranking system, commonly known as the "Pablo" rankings. Pablo rankings are currently published for college level volleyball, including men's and women's volleyball at the Division 1, 2, 3 and NAIA levels. Details of the ranking system will be provided, including the underlying (probabilistic) model and the key factors included in the assessment. Simulation results are used to explore the relationship between points scored and match outcomes. It will be shown that there is a simple relationship between point probability and win probability, and that this relationship gives results that accurately mirror real results, when accounting for the quality of the schedule. Performance of the system in predicting the outcomes of future matches will also be discussed.

Using Match Statistics To Assist in Grading Tennis Competitions

✱Neil T. Diamond, Monash University, Clayton campus, Melbourne, International 3041 Australia, *neil.diamond@buseco.monash.edu.au*

Key Words: Bradley-Terry model, Generlised Linear Models, Hierarchical Models, Random Effects

Most competition tennis players play in graded competitions, playing against players of hopefully similar ability. One of the most complex tasks of the organizing association is setting up the grades for each tennis season. For example, the North Suburban Junior Tennis Association, based in Melbourne, organizes two seasons per year, involving over 1500 players, 300 teams, and 40 grades each season. Getting the grading right is important to the players. Being graded in either too high or too low a grade can affect the motivation of players, their self-esteem, and enjoyment of the game. This talk describes an analysis of the last four seasons' match results using a modified Bradley-Terry model. Apart from improving the grading process, the model allows tracking of the improvement of individual players over time, as well as separate performance measures for singles and doubles.

Statistical Methods for Determining Optimal Rifle Cartridge Dimensions

✱Steven M. Anderson, Metropolitan State College of Denver, 2499 West 111th Place, Dept of Mathematics, Westminster, CO 80234, *sander69@mscd.edu*; Shahar Boneh, Metropolitan State College of Denver; Nels Grevstad, Metropolitan State College of Denver

Key Words: anova, factorial design, marksmanship, rifle

We have designed and carried out a statistical study to determine the optimal cartridge dimensions for a Savage 10FLP law enforcement grade rifle. Optimal performance is defined as minimal distance from the target's bull's-eye, and minimal group diameter. A full factorial block design with two main factors and one blocking factor was used. The two main factors were bullet seating depth and powder charge. The experimental units were individual shots taken from a bench-rest position and fired into separate targets. Additionally, 13 covariates describing various cartridge dimensions were recorded. The data were analyzed by a combination of methods, including ANOVA, Analysis of Covariance (ANCOVA), multiple regression, and response surface analysis. We will describe the experiment, the data analysis, and the interesting results.

Visualizing Managerial Strategies in Baseball

✱Steve C. Wang, Swarthmore College, Dept of Math and Stat, 500 College Ave, Swarthmore, PA 19081, *scwang@swarthmore.edu*

Key Words: exploratory data analysis, graphics, cluster analysis, sports, multivariate analysis

Baseball fans intuitively know which individual players are similar or different. For example, Rickey Henderson and Tim Raines are similar to each other, whereas both differ from Mike Schmidt. But which managers are similar, and which are unusual, in terms of the strategies they employ? Is Joe Torre more similar to Terry Francona or to Bobby Cox? Which managers use their bullpen in unorthodox ways? Which managers play “little ball” most often? We address questions such as these by exploring data on managerial decisions using multivariate analyses and high-dimensional visualization.

Is There Increased Parity in the NFL (Has Paul Tagliabue’s Dream Come True)?

✱Joseph Koopmeiners, University of Washington, Department of Biostatistics, Seattle, WA 98195, *jskoopm@u.washington.edu*

Key Words: football, state-space model, Bayesian

The St. Louis Rams and New England Patriots completed worst-to-first turnarounds to win Super Bowls XXXIV and XXXVI, respectively, leading football fans and sports writers to speculate that the NFL had entered a new era of parity. We investigate this hypothesis by modeling year-to-year autocorrelation with a Bayesian state-space model using NFL regular season results from 1970 to 2006. By transforming the covariance parameters we can model the variance and autocorrelation parameters in a regression context allowing for a simple comparison over time. Autocorrelation decreased steadily over time while the variance decreased from the 1970s to the 1980s but remained constant thereafter. This suggests that there is increased parity in the NFL but that it has manifested itself in decreased year-to-year autocorrelation rather than decreased variance in team strengths.

Assessing the Accuracy of Sports Predictions

✱Paul Stephenson, Grand Valley State University, Department of Statistics, 1133 MAK, 1 Campus Drive, Allendale, MI 49401, *stephenp@gvsu.edu*; John Gabrosek, Grand Valley State University; Dan Frobish, Grand Valley State University

Key Words: Sports prediction, Correlation, Agreement

The success of ESPN, sports periodicals and local sports radio stations has dramatically increased the number of “experts” that present their predictions for games, tournaments and seasons. Further, the fact that these proclaimed experts present their speculations on television, radio and the internet legitimizes their opinions. But to what degree has the accuracy of these experts been evaluated and what metrics are appropriate to evaluate their accuracy. In this presentation the authors will discuss several metrics for quantifying the accuracy of predictions. In addition, we will demonstrate the application of these metrics on a variety of sports, and we will use these metrics to compare the accuracy of several experts.

231 Modeling Mortality ●

Section on Statistics in Epidemiology, Section on Government Statistics, Social Statistics Section, Biometrics Section

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Reclassification of Socioeconomic Classes for Modeling Mortality from Leading Causes

✱Jay H. Kim, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *jhk0@cdc.gov*; Rong Wei, National Center for Health Statistics

Key Words: Mortality, Socioeconomic Variables, Principal Component

Instead of constructing socio-economic status (SES) groups using all 14 variables from the 2000 US Census that are considered as measures of social and economic status, we selected the variables that are highly correlated each other and to mortality to redefine SES groups. The redefined SES groups are also expected to be associated with mortality rates of five leading causes of death. Mortality rates for 805 health service areas are calculated from 2001–2005 mortality data file and linked to the SES variables from the census. SAS GLM and PRINCOMP procedures are used in selecting and constructing SES groups. To understand the relationship between mortality and the SES group, race-and-sex adjusted rates and race-and-sex specific rates are reviewed. The results are also compared to a previous study.

Mapping Empirical Bayes Estimates of Cancer Mortality Rates Adjusted for Smoking

✱Huilin Li, National Cancer Institute, 6120 Executive Boulevard, Executive Plaza South, Room 8089, Rockville, MD 20852-7242, *lih5@mail.nih.gov*; Barry I. Graubard, National Cancer Institute; Mitchell H. Gail, National Cancer Institute

Key Words: Empirical Bayes, SMR, cancer mortality rates, Poisson Gamma random effects Model, disease mapping, Health Service Areas (HSA)

The purpose of our research is to use information on known risk factors to produce adjusted cancer rates that might reveal other sources of spatial variation. We extend Clayton and Kaldor’s Poisson-Gamma random effects model to analyze cancer mortality rates with regression adjustment for smoking prevalence. In this way we can determine how much of the geographic variation can be explained by smoking and provide an estimate of the residual geographic variability after accounting for smoking. We compare conventional Standardized Mortality Ratio (SMR) estimates with SMR estimates adjusted for smoking and with empirical-Bayes SMR estimates from the Poisson-Gamma model that are also adjusted for smoking and are more stable for small areas. We analyze nationwide lung cancer and cardiovascular disease data in 805 geographic areas defined by Health Service Areas (HSA) level.

Identifiability of Bivariate Mixtures: An Application to Infant Mortality Models

✱Eric Y. Frimpong, University at Albany, Arts & Sciences Building, Room 237, 1400 Washington Avenue, Albany, NY 12222, *ef9999@albany.edu*; Timothy B. Gage, University at Albany; Howard Stratton, University at Albany

Key Words: Mixture models, Latent subgroups, Identifiability, infant mortality, birth weight

Identifiability of the parameters for a Mixture of Bivariate Densities (MBD) in the form $f(x,y;\beta,\theta,\pi)=\pi f(y|x;\beta_1)f(x;\theta_1)+(1-\pi) f(y|x;\beta_2)f(x;\theta_2)$ is considered. Particular attention is given to

$\theta_1\ne\theta_2$ (i.e. marginal of x is a nondegenerate mixture). Characterizations of identifiability that includes extensions of Hennig (2000), Hunter (2007) and Gage (2004) models are given. These identified models are applied to characterize latent subpopulations related to infant mortality and survival in several NCHS linked Birth/Death data sets.

Statistical Methods in Estimating Mortalities for Ages with Zero Observed Deaths

✱Rong Wei, National Center for Health Statistics, Room 3222, 3311 Toledo Rd., Hyattsville, MD 20782, *rrw5@cdc.gov*

Key Words: Mortality estimation, Life Tables, small area, Mixed Distribution, Two-part Model, Bayesian method

Annual age-specific mortality is a key estimate in constructing the US State-Specific Life Tables. The estimates are based on the observed number of deaths in each of the race-gender subpopulations. However the observed deaths are often zero for some young ages in small States during the data collection period. To estimate the mortality rates with "zero" deaths, several strategies have been applied: 1. taking an average of data from three consecutive years; 2. "borrowing" strengths from other States in the same data year using Bayesian methods; 3. LOESS smoothing with a two dimensional window of age and time; 4. applying probability models based on data from 30+ years; and 5. smoothing age mortality curves using Heligman-Pollard equation. In this paper we compare results from above methods applied on eight US States' data. A problem of "zero inflation" in estimation is also addressed.

Comparing Methods Decomposing Disparity in Life Expectancy by Disease

✱Charles Lin, U.S. Census Bureau, ACSD HQ-6H020A, Data Integration Division, 4600 Silver Hill Road, Washington, DC 20233, *charles.c.lin@census.gov*; Norman Johnson, U.S. Census Bureau

Key Words: Difference in Life Expectancy, Contribution of Diseases, Life-years Lost, Decomposition, Disparity

Lin & Johnson (*Statistics in Medicine* 2006) proposed a method of assessing the contribution of diseases (causes of death) to disparities in life expectancy among populations. Their methodology was derived by computing differences in life-years lost from diseases in the population, in contrast to other widely used methods (Arriaga 1984, 1989 Pollard 1988) which decomposed life expectancy differences by age group and then by disease. Note that a disease's contribution to the difference in life expectancy is restricted to the expected life-years, whereas the differences in life-years lost due to the disease are accumulated throughout all subsequent ages. A new method is also proposed for direct accessing the underlying contribution of diseases. Data from the National Longitudinal Mortality Study will be used to illustrate the differences in the estimates from these methods.

An Approach To Handle the Problem of Zeros Deaths in Estimating Mortality Rates

✱Anastasia Voulgaraki, National Center for Health Statistics, Room 3118, 3311 Toledo Rd., Hyattsville, MD 20782, *AVoulgaraki@cdc.gov*; Rong Wei, National Center for Health Statistics

Key Words: State-level mortality, mixed distribution, two-part models, hurdle models, zero-inflated models, bootstrap

In states with small subpopulations, the observed annual mortality rates are often zero, particularly among young ages. In estimating age-specific mortality, death rates are presented on logarithmic scales, and therefore zero rates are problematic. We use data from the NCHS public-use mortality files (1970-2002) to fit appropriate probability models and estimate expected values for replacing zero mortality. The probability models considered are based on mixed distribution and Poisson regression methods and are applied to mortality data from eight states with the problem of zero observed rates. The criterions used to compare the performance of the models are the root mean square error and the mean absolute error. The 95% confidence intervals are constructed using parametric bootstrap and asymptotic methods, where applicable. This paper describes the findings from our preliminary research.

On the Confidence Intervals of the Attributable Risk Under Cross-Sectional Sampling with Confounders

✱Md. Khairul Islam, Wayne State University, 110 E Warren Avenue, Detroit, MI 48201-1379, *islamk@med.wayne.edu*

Key Words: Attributable risk, Cross-sectional sampling, Confounder, Delta Method, Bootstrap, Confidence interval

The attributable risk is one of the most commonly used indices to measure the association between a risk factor and a disease outcome in public health and epidemiological studies. In this article, we have revisited the methods of constructing confidence interval estimators for attributable risk under cross-sectional sampling with confounders. A simpler expression has been derived using delta method and utilized in the construction of confidence intervals for attributable risk. A bootstrap version of interval estimation has also been considered as an alternative to the existing method which is straightforward and appropriate for constructing confidence interval. The performance of the bootstrap confidence interval has been justified with a real life example on the effect of smoking in infant death with age of mother and length of gestation as confounders.

232 Biomarkers, Sampling, and Measurement Issues

Section on Statistics and the Environment, Section on Statistics in Epidemiology

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Reconstruction of General Population Exposure to Perchloroethylene with Sparse Biomonitoring Data

✱Junshan Qiu, The University of Georgia, 101 Cedar Street, Department of Statistics, Athens, GA 30602, *qiuj@rx.uga.edu*

Key Words: Exposure reconstruction, Biomonitoring, PBPK, Bayesian

Physiologically based pharmacokinetics (PBPK) models have been extensively used to predict tissue dosimetry under controlled exposure conditions. Using biomonitoring data to reconstruct exposure profile with PBPK model is challenging. But the distributions of exposures or possible exposure pattern can be estimated based on the survey information of human activity diary or incomplete information of exposure. In this paper, we developed a method used to construct exposure concentration time profile of PCE for general population based on the biomonitoring data available. The distributions of exposure parameters are derived based on biomonitoring data or previous survey studies of human activity. Posterior distributions of model parameters were used to conduct Monte Carlo (MC) simulations. Bootstrap resampling was used to capture uncertainty of exposure conditions based on MC results.

Some Challenges Encountered in the Analysis of Biomarker Data

✱Stephen W. Looney, Medical College of Georgia, Department of Biostatistics, 1120 15th St., AE-3020, Augusta, GA 30912-4900, *slooney@mcg.edu*; Joseph L. Hagan, LSU Health Sciences Center

Key Words: normality, homoscedasticity, correlation coefficient, measures of agreement, t-test, McNemar test

In this presentation, we provide a description of some challenges that are often encountered in the analysis of biomarker data, and offer recommendations on how best to deal with those challenges. The statistical issues that we consider are illustrated with examples taken from the biomarker literature and include: assessment of distributional assumptions, assessment of homogeneity assumptions, comparison of correlation coefficients, and comparison of biomarkers. Our recommendations are based on published statistical evidence and we indicate how widely available statistical software can be used to carry out the recommended analyses.

Approximate Inferential Procedures Based on Samples with Nondetects

✱Avishek Mallick, University of Louisiana at Lafayette, Lafayette, LA 70504, *axm9357@louisiana.edu*; K. Krishnamoorthy, University of Louisiana at Lafayette; Thomas Mathew, University of Maryland Baltimore County

A generalized variable approach (GV) based on the maximum likelihood estimates for analyzing a sample with nondetects is proposed. The method is applied for the normal case, and is illustrated for constructing prediction limits, one-sided tolerance limits and for setting lower confidence limit for a survival probability. The validity of the methods is evaluated using Monte Carlo simulation. Simulation studies show that the GV approach is satisfactory as long as the proportion of nondetects is not too large. Approximate inferential methods are also given for a gamma distribution. Monte Carlo evaluation of the procedures indicates that the method works satisfactorily for constructing tolerance limits, prediction limits and estimating survival probability for a gamma distribution. The methods are illustrated using some practical examples.

A Risk-Based Approach to Setting Environmental Limits on Turbidity During Large-Scale Dredging Operations

✱David R. Fox, University of Melbourne, PO Box 4102, Parkville, International 3052 Australia, *david.fox@unimelb.edu.au*

Key Words: environmental protection, risk minimisation, seagrass protection, dredging, turbidity, control charting

Turbidity is a naturally occurring phenomenon associated with any water body. Human-induced turbidity through mechanical disturbance of sediments or direct deposition of particulate matter can have profound, negative environmental impact. Of particular concern is the potential for loss of large tracts of sea grass due to reduced light availability. Economic imperatives are driving large-scale dredging of shipping channels in Australia, New York and elsewhere around the world. Conventional approaches to assessing turbidity-related impacts have relied on fairly simplistic regression models to describe benthic light climates. In this paper we describe a new, risk-based approach to establishing limits on turbidity so as to meet minimum light requirements with some prescribed probability. The method has been incorporated into a real-time surveillance system.

Using a Bayesian Hierarchical Measurement Error Model To Estimate Individual Observer Detection Rates in Aerial Transect Surveys

✱Mark C. Otto, U.S. Fish and Wildlife Service, Patuxent Wildlife Research Center, 11510 American Holly Dr, Laurel, MD 20708-4002, *Mark_Otto@FWS.Gov*

Key Words: Poisson-Log Normal

We compare visibility corrections in the North American Breeding Waterfowl Survey using design-based and Bayesian hierarchical model estimators. The survey currently uses a combined ratio estimator to assess the proportion of ducks missed from the fixed wing counts compared to either ground or helicopter counts of the same segments. Counts are pooled over enough years and observer areas to provide stable estimates of detection. A Bayesian hierarchical model can estimate detection rates using a measurement error model that includes random individual observer effects. Since pilot and observer count adjacent sides of the same segments, all the data can be used to estimate observer difference rather than just the segments with ground counts. The hierarchical model yield more stable detection rates and population totals.

A Series of Experiments About Wireless Sensor Networks

✱Kristian Lum, Duke University, *kcl12@duke.edu*

Key Words: wireless sensor networks

Often the limiting factor in the quality of inference from wireless environmental sensor networks is due to the battery life of the nodes. Researcher must make decisions about how long the network's life span should be and how precise of estimates are desired for their particular application. My presentation shows several small experiments with the goal of looking at how inference on parameters degrades as samples are taken more sparsely and under different sampling schemes. It also looks at the trade-offs between parameter inference, missing data estimation, and conservation of battery (using fewer transmissions). Some of the models used in investigating these questions are a plain AR(1), a two dimensional AR(1), AR with drift, AR with shocks/other predictors, and a time-series of binary data.

233 Issues in Substance Abuse and Mental Health Research

Section on Health Policy Statistics
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Using a Yardstick Approach To Enrich Understanding of Treatment Effect on Urge To Smoke: An Application to a Smoking Cessation Trial

✱Joseph C. Cappelleri, Pfizer, Inc., 50 Pequot Avenue, MS 6025-B2257, New London, CT 06320, *joseph.c.cappelleri@pfizer.com*; Andrew G. Bushmakin, Pfizer, Inc.; Riad Dirani, Johnson & Johnson Pharmaceutical R&D, LLC

Key Words: effect size, clinical significance, patient-reported outcome, urge to smoke, smoking cessation, varenicline

We illustrate a way to enrich the interpretation of treatment effect on patient-reported outcomes. The study involved 1,025 smokers randomized to receive varenicline, bupropion, or placebo for 12 weeks. Urge to smoke, a patient-reported outcome, was measured by the Minnesota Nicotine

Withdrawal Scale (MNWS). We divided the treatment effect size (difference between treatments on urge to smoke averaged over the first 7 weeks of treatment) by the criterion effect size (difference between non-smokers and smokers on urge to smoke at week 12 in the placebo group). Compared with placebo, varenicline reduced urge to smoke on MNWS that was about six-tenths of the way found between non-smokers and smokers. By benchmarking against two distinct clinical groups, the interpretation of treatment effect on patient-reported outcomes can be enriched.

Alcohol Outlet Density and Alcohol Consumption in Los Angeles County and Southern Louisiana

✱Matthias Schonlau, RAND Corporation, 4570 Fifth Ave, Suite 600, Pittsburgh, PA 15213, *matt@rand.org*; Deborah Cohen, RAND Corporation; Richard Scribner, Louisiana State University Health Sciences Center; Thomas A. Farley, Tulane University School of Public Health and Tropical Medicine; Katherine Theall, Louisiana State University Health Sciences Center; Ricky N. Bluthenthal, California State University; Molly Scott, RAND Corporation

Key Words: acohol consumption, alcohol outlets

We assessed the relationship between alcohol availability, as measured by the density of off-premise alcohol outlets, and alcohol consumption in Los Angeles county and southern Louisiana. Consumption information was collected through a telephone survey of 2881 households in Los Angeles County and pre-Katrina southern Louisiana nested within 220 census tracts. Respondents' addresses were geocoded and both neighborhood (census tracts and buffers of varying sizes) and individual (distance to the closest alcohol outlet) estimates of off-sale alcohol outlet density were computed. Alcohol outlet density was not associated with the percentage of people who were drinkers in either site. Alcohol outlet density was associated with the quantity of consumption among drinkers in Louisiana but not in Los Angeles.

Health Consequences of Alcohol Abuse: An Ecological Model

✱Yasmin H. Said, George Mason University, 23044 Winged Elm Dr., Clarksburg, MD 20871, *ysaid99@hotmail.com*

Key Words: Stochastic hierarchical model, Dirichlet distribution, Stochastic directed graph, acute outcome, HIV, time-of-day

The users of alcohol and the environment in which they are embedded may be thought of as analogous to an ecological system. Controlling specific aspects of this system (e.g., drinking and driving) may result in unintended consequences (e.g., an increase in domestic violence as the result of high volume drinkers spending more time at home due to the emphasis on policing of drunk driving program). Another example could be promiscuous behavior as a result of drinking too much, which, in turn, might lead to infection with sexually transmitted diseases (STD) and Human Immunodeficiency Virus (HIV). In this paper we develop a model of this system and investigate interventions for the simultaneous suppression of negative health consequences.

An EM Algorithm for Zero-Inflated Negative Binomial Regression Based on a Poisson Mixture Representation

✱Chih-Nan Chen, Cambridge Health Alliance, Center for Multicultural Mental Health, 120 Beacon Street, 4th Floor, Somerville, MA 02143, *chihnan@gmail.com*; Xiao-Li Meng, Harvard University

Key Words: Zero-Inflated negative binomial regression model, Overdispersion, Count data with excess zeros, EM algorithm

Zero-inflated negative binomial regression model is quite useful for count data with extra zeros and over-dispersion. Fitting such a model, however, is not a trivial task, particularly when the sample sizes are small. We found that an EM algorithm based on a mixture Poisson representation of the negative binomial density can help to reduce some instability of existing algorithms, because the new implementation takes advantage of the well-behaved Poisson regression fitting in its M step. Application is given to the analysis of visit numbers in the service use of mental health care for the National Latino and Asian American Study (NLAAS), a problem that originally motivated us to develop this algorithm because of the instability problems we encountered with several existing algorithms.

Utilizing Stakeholders To Inform Mental Health Treatment Practices

✱Christopher J. McKinney, Mental Health Center of Denver, 4141 E. Dickenson Pl., Denver, CO 80222, *christopher.mckinney@unco.edu*; Kathryn DeRoche, Mental Health Center of Denver; Antonio Olmos, Mental Health Center of Denver; Mike Hester, Mental Health Center of Denver; Donald Sammons, Mental Health Center of Denver

Key Words: Measurement, Mental Health, Stakeholders, Quality Control, Mixed Methods, Evaluation

Incorporating and representing consumer, clinician, and additional stakeholder input into mental health outcomes can be a difficult task, but is necessary for optimal development, analysis, and dissemination of findings. In order to overcome these difficulties MHCD has implemented several policies to ensure a holistic development of recovery outcome measures by incorporating multiple stakeholders. The incorporation of stakeholders' input early on in the measurement development process and continuing through implementation of the tools allows many issues to be easily resolved, including internal and external validity, and reliability. The current paper will review this process of incorporating stakeholders in all areas of outcomes evaluations, including issues that can make the process much easier to implement system-wide.

Risk Factors Associated with Injection Initiation Among Drug Users: Comparing Hierarchical Modeling with Traditional Logistic Regression

✱Shahid Kamal, University of the Punjab, College of Statistics and Acturial Sciences, University of the Punjab, Lahore, International 56000 Pakistan, *kamal_shahid@yahoo.com*; Muhammad A. Raza, GC University; Rehan A. Khan, PU

Key Words: Multilevel modelling, logistic regression, IDU, DU

In the recent past there has been a significant shift from traditional modes of drug use (chasing, sniffing, and smoking) to injecting drugs in Pakistan, which has also significantly contributed to increasing epidemics of HIV/AIDS, hepatitis B and C virus among this high risk group. The purpose of the study is to evaluate the extent to which demographical, sexual behavior and drug use practices predict initiation of injecting drug use in Pakistan. The study group consisted of 457 male drug users, 257 of whom were injecting drug users in contrast to the 200 drug users who were not injecting drug user. Both descriptive and inferential studies have been made to explore the important features of the study. Conventional logistic and multilevel logistic regression was employed to identify the independent effect from potential risk factors of transition into injection.

234 The Decennial Census: Past and Future

Section on Government Statistics, Social Statistics Section

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

An Overview of the 2010 Census Design

✱Frank Vitrano, U.S. Census Bureau, 4600 Silver Hill Road, Room 3H174, Suitland, MD 20746, *frank.a.vitrano@census.gov*

Key Words: 2010 Census, integration, management

This paper will provide a description of the key components of the 2010 Census design. It will include a brief discussion of the vision, goals, and objectives of the 2010 Census; a discussion of the six functional areas of the census and the corresponding operations and systems that complete the function; and, a description of the key management structures and strategies for ensuring a well-defined and successful census design and program. These management structures include centralizing fundamental integration activities such as requirements, budget, risk, and schedule management in the Decennial Management Division as well as forming interdivisional teams such as the Census Integration Group and the Decennial Systems Architecture Team that will support operational and technical integration.

Reducing Risks in the 2010 Census

✱Sally Obenski, U.S. Census Bureau, 4600 Silver Hill Road, 3H161, Suitland, MD 20746, *sally.m.obenski@census.gov*; James Treat, U.S. Census Bureau; Dan Weinberg, U.S. Census Bureau

Key Words: risk, integration, plans, 2010 Census, assessments

This paper focuses on strategic approaches to managing risk in the 2010 Census. It expands on the management structures discussed in the overview paper, with an emphasis on risk reduction. It will discuss how the Decennial Management Division has implemented program-wide integration activities for managing program risk. It describes how the Census Bureau is managing information technology risks by using modeling and simulation and integrated system testing. It further discusses the use of highly skilled contractors to conduct system analyses and assessments of critical program components. Finally, it describes the formal risk management process that the Census Bureau is using along with a few examples of specific risk plans currently in place.

Ensuring That Processes and Information Inputs and Outputs Are Accurate: Operational Integration

✱James Treat, U.S. Census Bureau, 4600 Silver Hill Road, 3H177, Suitland, MD 20746, *james.b.treat@census.gov*; Frank Vitrano, U.S. Census Bureau

Key Words: operations, inputs, 2010 Census, outputs

This paper focuses on the need for and strategies used to ensure that the defined operations comprising the 2010 Census are integrated and are passing the right data to the right operation in the right sequence. Included in this discussion is the move to an analytic Operations and Systems Plan that focuses on inputs and outputs; the development of an operational node view; the identification and definition of "hubs" that form the critical data manipulation points of the design; and, the modeling and simulation that has been conducted to test throughput and effects of schedule delays. Additionally, this paper discusses the implementation of key processes for change, issues, and schedule management.

Effectively Communicating a Complex Program

✱James Dinwiddie, U.S. Census Bureau, 4600 Silver Hill Road, 3H171, Suitland, MD 20746, *james.l.dinwiddie@census.gov*; Frank Vitrano, U.S. Census Bureau

Key Words: 2010 Census, oversight, communications, external

This paper focuses on the importance, challenges, and strategies of communicating the 2010 Census' plans, decisions, and issues to a wide array of internal and external stakeholders. It discusses how program-related communications flows up, down, and across organizational structures. It discusses our approaches to massive requests for information and interviews by a number of oversight entities, including the Government Accountability Office, the Department of Commerce, and the Office of Management and Budget. It also includes our strategies for achieving consistency in a highly time-sensitive and many times volatile program development and implementation environment, given the needs to respond to external decisions (e.g., Continuing Budget Resolutions, additional security measures).

Estimating the Average Number of Persons Per Household: A Regression Approach

✱Kirsten K. West, U.S. Census Bureau, Population Division, Washington, DC 20230, *kirsten.k.west@census.gov*; Antonio Bruce, U.S. Census Bureau

Key Words: estimation, regression, population estimates

The Population Division of the U.S. Census Bureau recently started a project to explore strategies for producing county level housing unit based population estimates. The housing unit method relies on the number of occupied housing units and the average number of persons per household (PPH) at a particular time. Research by Smith, Nogle, and Cody (Demography, November, 2002) confirmed that regression models based on symptomatic indicators of PPH change can produce more precise and less biased county PPH estimates than methods that rely on the most recent decennial census for estimation. Following their work, we propose to use regression models to develop county level PPH for postcensal years and then compare them against PPH directly derived from the American Community Survey.

Convergent Multivariate Graphics for Capturing the Naming Patterns in U.S. Census Data Over Two Centuries

✱Whitney Erickson, Brigham Young University, 551 North 200 East #205, Provo, UT 84606, *we13sports@gmail.com*; Angel Venegas, Brigham Young University; Alexandra Brattos, Brigham Young University; Malvina Salash, Birgham Young University; Donald Walker, Brigham Young University; Jessica H. Scott, Brigham Young University; Bruce L. Brown, Brigham Young University

Key Words: census data, personal names, onomastics, multivariate graphics

Male and female given names data (from the United States Census database and the Social Security database) are available from 1801 to the present. These data present a rich and complex source for studies of naming patterns over time. A number of multivariate graphical data analysis methods are demonstrated for capturing the holistic patterns of changes in naming practices over time, including: profiles of box and whiskers plots of distributions of name frequencies over successive decades, bivariate and trivariate scatterplots with box and whisker marginal plots comparing transition over two or three decades, linked principal components plots and dendrograms from cluster analysis to capture holistic patterns, etc.

235 Bayesian Applications and Methods in Biology ●▲

Section on Bayesian Statistical Science
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Long-Term HIV Dynamic Models Incorporating Drug Adherence and Resistance for Prediction of Virologic Responses

✱Yangxin Huang, University of South Florida, College of Public Health, MDC 56, 13201 Bruce B. Downs Blvd., Tampa, FL 33612, *yhuang@health.usf.edu*

Key Words: Adherence, Bayesian nonlinear mixed-effects models, time-varying treatment efficacy, deviance information criterion, long-term HIV dynamics, longitudinal data

Imperfect adherence and drug resistance (DR) to prescribed antiretroviral (ARV) therapies are important factors explaining the resurgence of virus. A better understanding of the factors responsible for the virologic failure is critical for the development of new treatment strategies. We here develop a mechanism-based reparameterized differential equation models with incorporating adherence (MEMS or questionnaires) and DR for characterizing long-term viral dynamics with ARV therapy. A Bayesian NLME modeling approach is investigated for estimating parameters and comparing effects of different adherence assessments based on an AIDS trial dataset. The results indicate that the drug adherence combined with confounding factor, DR significantly predicts virologic responses (VR). Our study suggests that our models are effective in establishing a relationship of VR with drug adherence and DR.

Bayesian Mixture Modeling for Interval-Censored Age Onset of Puberty

Bin Huang, Cincinnati Children's Hospital Medical Center; ✱Lili Ding, Cincinnati Children's Hospital Medical Center, Center for Epi/Bio, MLC 5041, 3333 Burnet Ave, Cincinnati, OH 45229, *lili.Ding@cchmc.org*; Siva Sivaganisan, University Cincinnati

Key Words: Bayesian, Mixture Model, Interval Censor, Case Study, Puberty

Age onset of puberty marks the time point when adolescent start developing secondary sexual characteristics. Adolescents who experience onset of puberty at earlier or later age are at risk for adverse physical and psychosocial development. Since subject could only seen in a predetermined time interval, this is typically an interval censored measure. Current researches use an arbitrary classification rule by assigning adolescent whose age onset of puberty fell below 20% or above 80%-tile as early or later timer. We propose using a flexible mixture Bayesian modeling approach for classifying adolescent according to their interval censored age onset of puberty. The number of classes is determined based on a cross validation index. The individuals are classified into one of the classes based on their posterior probabilities. Its performance is illustrated using simulated and real data.

Bayesian Methods for Models of Virus Production by Infected Cells

✱Cavan Reilly, Universty of Minnesota, 420 Delaware St SE, A460 Mayo Bldg MMC 303, Minneapolis, MN 55403, *cavanr@biostat.umn.edu*

Key Words: Bayesian models, prior information, HIV/AIDS

In the early stages of SIV infections, productively infected resting memory CD4+ T cells comprise more than 90 percent of the productively infected cell population. While in snap-shots, the infected resting cells apparently produce much less virus than infected activated CD4+ T cells, the resting cells produce virus for longer periods of time. Here we develop models to account for these and other factors impacting the quantity of virus produced by a cell, and use Bayesian methods to conduct inference for this class of models. We are able to show that a substantial proportion of the virus that is measured in infected animals comes from the resting cells using only counts of infected cells surrounding samples drawn from each cell population and other estimates available in the literature.

Detection of Task-Related Activation in fMRI Data with Segmentation Information from Structural MRI

✱Namhee Kim, The Ohio State University, Department of Statistics,, 1958 Neil Avenue, Columbus, OH 43210, *namhee@stat.ohio-state.edu*; Prem K. Goel, The Ohio State University; David Q. Beversdorf, University of Missouri

Key Words: Bayesian model, combining information

As the intensity levels in Magnetic Resonance Image (MRI) data differ by tissue type, image segmentation results have been used more for clinical purposes to investigate morphological changes in brain tissue. Segmentation procedures can produce structural MRI image partitions by relevant tissue types: grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF), and background classes. Furthermore, as task-related true activations in the brain in functional MRI (fMRI) data are expected to be mostly on grey matter tissue, use of segmentation information is expected to produce more reliable activation maps in fMRI context. Recently, the adaptation of activation maps referring to the segmentation results has been attempted. In this study, we propose a method for the detection of activation combining information from segmentation results in MRI within the Bayesian framework.

Network-Based Auto-Probit Modeling for Protein Function Prediction

✱Xiaoyu Jiang, Boston University, 111 Cummington Street, Boston, MA 02215, *xiaoyu@math.bu.edu*; Eric Kolaczyk, Boston University; David Gold, Boston University

Key Words: Protein function prediction, Conditional autoregressive, Network

Predicting the functional roles of proteins based on various genome-wide data, such as protein-protein interaction networks (PPI), has become a canonical problem in high-throughput computational biology. Approaching the problem as binary classification, we propose a network-based extension of the spatial auto-probit model. In particular, we develop a fully Bayesian probit-based framework, with a latent multivariate conditional autoregressive (CAR) Gaussian process, where the latter encodes a measure of protein functional similarity influenced by network topology. We use this framework to predict protein functions, for function defined in the Gene Ontology (GO) database, a popular rigorous vocabulary for biological functionality. Markov Chain Monte Carlo methods are used to gain posterior estimates. A cross-validation study is performed on the data from the yeast Saccharomyces cerevisiae.

A Spike and Slab Centering Distribution in Dirichlet Process Mixture Models for Gene Expression

✱Sinae Kim, The University of Michigan, Department of Biostatistics, 1420 Wasington Heights, Ann Arbor, MI 48109, *sinae@umich.edu*; David B. Dahl, Texas A&M University; Marina Vannucci, Rice University

Key Words: Dirichlet process mixture models, Gene expression data, Multiple testing, Spike and slab prior

Model-based clustering methods using Dirichlet process (DP) mixture models have been proposed to exploit clustering for increased sensitivity in multiple hypothesis testing. Rather than yielding a probability of a hypothesis for each object, existing methods can only provide a ranking of the objects by their evidence for a particular hypothesis. In this work, we adapt the framework on Dahl et al. (2008) to accommodate point null hypotheses. For that, we use a spike and slab distribution which is a mixture of both a point-mass distribution and a continuous distribution as the centering distribution for the Dirichlet process prior. The method yields probabilities that genes follow the hypotheses of interest, whether those hypotheses be sharp or not. We apply our method in gene expression context and show how to simultaneously infer gene clustering and differential gene expression.

236 Applications of Survival Analysis ●

ENAR, WNAR, Section on Statistics in Epidemiology, Biometrics Section

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Gray's Time-Varying Coefficient Model with Censored Covariates

✱Gina D'Angelo, Washington University School of Medicine, 660 South Euclid Avenue, Box 8067, Division of Biostatistics, St. Louis, MO 63110-1093, *gina@wubios.wustl.edu*; Lisa Weissfeld, University of Pittsburgh; Chung-Chou Chang, University of Pittsburgh

Key Words: biomarkers, censored covariates, survival, time-varying coefficients, weighted pseudolikelihood

While Gray's time-varying coefficient model (1992) can be used to explore nonproportional hazards and to handle data when the outcome is censored, it cannot be used to handle data when both the outcome and the covariates are censored. In the example that we consider here, biomarkers for sepsis are often censored because bioassays are unable to accurately detect marker levels that are below a given threshold. We propose a new method that can be applied to handle censored covariates in Gray's time-varying coefficient model. In this new method, we construct a weighted penalized pseudolikelihood that incorporates inverse probability weighting for censored covariates. In the example of sepsis, we compare the results when we use this new method with the results when we fill in the censored values with lower threshold values. The GenIMS investigators are contributing authors for this abstract.

Sensitivity Analysis To Investigate the Impact of a Missing Covariate on Survival Analyses Using Cancer Registry Data

✱Brian L. Egleston, Fox Chase Cancer Center, Biostatistics Facility, 333 Cottman Avenue, Philadelphia, PA 19111-2497, *Brian.Egleston@fccc.edu*; Yu-Ning Wong, Fox Chase Cancer Center

Key Words: SEER, Cancer, missing data, proportional hazards regression, sensitivity analysis

Having substantial missing data is a common problem in cancer registry data. We propose a sensitivity analysis to evaluate the impact of data that is potentially missing not at random in survival analyses using Cox proportional hazards regressions. We apply the method to an investigation of how tumor grade moderates the effect of surgery on mortality outcomes in patients with metastatic renal cell carcinoma in the NCI's Surveillance Epidemiology and End Results (SEER) registry. Tumor grade is an important component of risk stratification for patients with kidney cancer. However, many individuals in SEER are missing tumor grade information. In this observational study, we found that removal of the primary tumor was associated with improved survival, but that the magnitude of the effect depended on assumptions about the relationship of grade with missingness.

Estimating Predictive Error for Left- and Right-Censored Survival Data with Time-Varying Covariates

✱William R. Prucka, The University of Alabama at Birmingham, 2013 Highlands Dr., Hoover, AL 35244, *prucka@uab.edu*; Christopher S. Coffey, The University of Alabama at Birmingham

Key Words: Survival, Prediction Error, Resampling, Bootstrap, Cross validation, IMLPlus

In clinical practice it is beneficial to develop prognostic equations that can delineate at-risk patients using baseline and time-varying data. For outcomes such as survival, predictive models must be estimated with incomplete follow-up, or right censoring. Additionally, patients may not initiate therapy at disease incidence, resulting in left censoring. We extend a published resampling procedure to estimate prediction error for right and left censored data with time varying covariates. The method assesses an arbitrary prediction rule via K-fold cross validation or the 0.632+ bootstrap estimator. Implemented in IMLPlus, a free extension to SAS, our software provides intuitive model fit, selection, and visualization. We evaluate the method through simulation to investigate the impact of left censoring, hazard shape, sample size, and time varying covariates on prediction error.

Using a Point System Table To Improve the Communication of Statistical Estimates of Risk to Clinicians and Patients

✱Darcy Hille, Merck & Co., Inc., Mailstop UG1C-46, 351 Sumneytown Pike, North Wales, PA 19454, *darcy_hille@merck.com*

Key Words: Cox proportional hazards model, points system, risk estimates, statistical communication

A Cox proportional hazards model identified baseline variables that had a significant impact on the occurrence of the primary endpoint in a large randomized clinical trial to prevent cardiovascular events. A point table was developed which assigned points to various levels of the risk factors, including an interaction term, by adapting the point system algorithm developed by Sullivan (2004). Risk estimates defined by point totals were compared to risk estimates defined by the Cox model; the weighted kappa for agreement was 0.89. The point table represents an intuitive method of quantifying the risk of the primary endpoint, and is a useful tool for communicating a complicated statistical model to clinicians and patients.

A Comparison Between Rank Regression and Maximum Likelihood Methods of Estimation for Weibull Regression

✱Laura J. Freeman, Virginia Polytechnic Institute and State University, 213-A Hutcheson Hall, Blacksburg, VA 24061, *lamorgan@vt.edu*; Denisa A. Olteanu, Virginia Polytechnic Institute and State University; Geoffrey Vining, Virginia Polytechnic Institute and State University

Key Words: Reliability, Weibull Regression, Rank Regression, Maximum Likelihood

Weibull regression analysis is often an appropriate technique for analyzing reliability data. Time to failure is well modeled by a 2-parameter Weibull distribution. The maximum likelihood method of estimation is a popular method of estimating regression parameters due to its asymptotic properties. These properties however hold when sample sizes are large which is not common in practice for reliability data, often a model is desired after only a very small number of failures. Rank-regression has been used by practitioners because of its simplicity and ease of understanding. These two methods are compared for small (10), medium (100) and large (1000) sample sizes with varying numbers of failures. Four different failure modes, infinite, random, wear out and rapid wear out are evaluated under the two methods of estimation. Recommendations are made based on simulation results.

Mixture Hazard and Change-Point Models in Survival Analysis: An Application to Predicting Suicide Attempts

✱Hanga Galfalvy, Columbia University, Department of Psychiatry, 1051 Riverside Drive, Box 42, New York, NY 10032, *hcg2002@columbia.edu*; Maria A. Oquendo, Columbia University; John J. Mann, Columbia University

Key Words: survival analysis, change-point, mixture hazard

For prognostic models in medicine with multimodal or bathtub-shaped hazard function, 2 questions need to be answered: 1) do(es) the change-point(s) exist, 2) if so, a) does it reflect the existence of a mixture population, with different hazard functions for subgroups of subjects, or b) is it part of the natural course of the disease. For simple forms of the hazard plot, testing for a constant/linear functional form vs. an alternative of piecewise linear hazard can answer the first question. Next, survival regression models can test for patient subgroups with hazard functions of different forms, although these will not necessarily align directly with the change-points identified. We will illustrate a combination of statistical methods that test both hypotheses on a prognostic model of risk of suicide attempt for subjects in a major depressive episode.

Strong Consistency of the Intercept Estimator in the Semiparametric Accelerated Failure Time Model

✱Ying Ding, The University of Michigan, 1980 Traver Rd. Apt 211, Department of Biostatistics, Ann Arbor, MI 48105, *yingding@umich.edu*; Bin Nan, The University of Michigan

Key Words: Accelerated failure time model, censored survival data, unbounded covariate support, Gehan weights

A long existing problem in linear regression for censored survival data is the estimation of intercept when error distribution is unspecified. This becomes particularly important when prediction of failure times for new observations is of major interest. We show that the intercept can be consistently estimated when supports of some covariates are not restricted to finite intervals, provided consistent estimates of slopes. The intuition is that with unbounded covariate support, even the follow-up time is finite, supports for the censoring time and the survival time on the residual scale are equivalent, which satisfies a sufficient condition in Susarla & Van Ryzin (1980) and Stute & Wang (1993). We also show that, without assuming bounded covariate support, slope estimators obtained by solving the rank-based estimating equation using Gehan weights are consistent and asymptotically normal.

237 Stochastic Modeling of Biological Data

Biometrics Section

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Branching Process Models of the Cell Cycle

✱Peter Olofsson, Trinity University, *polofsso@trinity.edu*

Key Words: branching process, cell cycle, stochastic process, population dynamics, mathematical biology

A class of stochastic models for the cell cycle is proposed. Specifically, we propose a general branching process where the phases of the cell cycle are recorded by the use of so called random characteristics. As illustrations, we consider cell populations with quiescence and cell populations with S-phase asynchronicity. Both of these have previously been approached with deterministic differential-equation based models to which we compare our approach and results.

Matching Errors in Closed-Population Capture-Recapture Experiments Based on DNA Sampling

✱Ross M. Gosky, Appalachian State University, 120 Buttercup Court, Boone, NC 28607, *goskyrm@appstate.edu*; Leonard Stefanski, North Carolina State University

Key Words: Capture-Recapture, Matching Errors, DNA sampling

Capture-Recapture models are used to estimate the unknown sizes of animal populations. When DNA samples, such as hair samples, are used to identify the capture of a particular animal, the possibility for matching errors exists across capture periods. We focus particularly on cases where two closely related animals may have similar DNA structures, and their DNA profiles may not be uniquely identifiable. We term this type of error a Shadow Match error. Such errors cause negative bias in estimates of population size. We introduce a statistical model accounting for these matching errors, and through simulation we show that these models produce more accurate estimates of population size when some information about the population structure can be introduced into the study.

Modeling Risk Factors for Alzheimer's Disease Progression Using a Nonhomogeneous Markov Process

✱Rebecca A. Hubbard, University of Washington, 4311 11th Ave NE, Suite 300, Seattle, WA 98105, *rebecca.hubbard@gmail.com*; Andrew Zhou, University of Washington

Key Words: longitudinal, Markov process, Alzheimer's disease, disease progression, multistate disease, panel data

Identifying individuals at risk of developing Alzheimer's disease (AD) is important for understanding the natural history of the disease and most effectively targeting interventions. One group with high probability of progression to AD are subjects suffering from mild cognitive impairment (MCI). We propose a non-homogeneous Markov process model to characterize transitions between disease states defined by normal cognition, MCI, AD, and death and identify risk factors for conversion. Nonhomogeneous Markov process models are particularly useful in the case of AD because cognitive status is ascertained only at periodic follow-up visits and conversion rates are known to be highly age dependent. We apply this model to the Uniform Data Set, a longitudinal study of subjects evaluated at one of the National Institute on Aging's Alzheimer's Disease Centers.

Estimating the Number of Species with Quadrats Sampled Without Replacement

✱Tsung-Jen Shen, National Chung Hsing University, 250 Kuo Kuang Road, Tai-Chung, Taiwan., (Department of Applied Mathematics), Tai-Chung, Taiwan, *tjshen@amath.nchu.edu.tw*; Fangliang He, University of Alberta

Key Words: quadrat sampling, richness estimator, sampling without replacement, zero-truncated binomial distribution, maximum likelihood estimate, modified beta distribution

Most richness estimators currently in use are derived from models that consider sample with replacement or from the assumption of infinite populations. Neither of the assumptions is suitable for sampling sessile organisms such as plants where quadrats are often sampled without replacement and the area of study is always limited. In this talk we propose an incidence-based parametric richness estimator that considers quadrat sample without replacement in a fixed area. The estimator is derived from a zero-truncated binomial distribution for the number of quadrats containing a given species and a modified beta distribution for the probability of presence/absence of a species in a quadrat. The maximum likelihood estimate of richness is explicitly given and can be easily solved. The performance of the estimator is tested using two tree data sets where the true numbers of species are known.

A Time-Dependent Poisson Random Field Model for Polymorphism Within and Between Two Related Biological Species

✱Amei Amei, Univeristy of Nevada Las Vegas, Department of Mathematical Sciences, University of Nevada Las Vegas, Las Vegas, NV 89154, *amei.amei@unlv.edu*; Stanley Sawyer, Washington University in St.Louis

Key Words: Poisson random field, diffusion approximation, population genetics, mutation, aligned DNA sequences

We derive a time-dependent Poisson random field to model genetic differences within and between two species that share a relatively recent common ancestor. We first consider a random field of Markov chains that describes the fate of a set of individual mutations. This field is then approximated by a time-dependent Poisson random field from which we can make inferences about the amounts of mutation and selection that have occurred in the history of observed aligned DNA sequences.

Stochastic Modeling for Blood Glucose and Insulin Dose in a Simulation Study of Glucose Control Strategies in the Adult ICU

✱Hui Zheng, Harvard Medical School, 50 Staniford Street Suite 560, Boston, MA 02114, *hzheng1@partners.org*; David Schoenfeld, Massachusetts General Hospital; Taylor Thompson, Massachusetts General Hospital

Key Words: monte carlo simulation, Bayesian model

To control patients' blood glucose level in the ICU, glucose is regularly monitored and insulin infusion rate adjusted accordingly. We use a statistical model to analyze the relationship between insulin dose and glucose and conduct computer simulations of the effect of different treatment strategies. The statistical model assumes the error variance in glucose change rate is proportional to the length of time interval between successive monitoring. We use longitudinal clinical data to obtain model parameters and validated our main model assumptions. We vary simulated monitoring frequency and assess the potential benefits of an Bayesian insulin dose algorithm that helps reduce variability in glucose. Simulations suggest that increased monitoring frequency can lead to substantially better glucose control. Sensitivity analyses suggest our findings are insensitive to model assumptions.

Modeling of PSA Dynamics of Prostate Cancer Patients Receiving Neoadjuvant Hormone Therapy Prior to Radiation Therapy

✱Mariel S. Lavieri, The University of British Columbia, 2053 Main Mall, Vancouver, BC V6T 1Z2 Canada, *Mariel.Lavieri@sauder.ubc.ca*; Martin L. Puterman, Sauder School of Business; Scott Tyldesley, British Columbia Cancer Agency ; Steven Shechter, The University of British Columbia ; Tom Pickles, British Columbia Cancer Agency

Key Words: prostate cancer, medical decision making, Kalman filter, disease modeling, Bayesian updating

This talk describes the use of the Kalman filter to update knowledge about PSA kinetics of high-intermediate and high risk prostate cancer patients that receive neoadjuvant hormone therapy prior to radiation therapy. Initial believes are based on population characteristics. Estimates of the nadir (which might be linked to maximal tumor regression) are updated as new information becomes available. The model could be used as input in the monitoring and treatment decision making process of prostate cancer patients.

238 Risk Analysis

Business and Economics Statistics Section, Section on Risk Analysis

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Empirical Relationships Among Various Statistics Measuring Discriminatory Power

✱John Zuo, Wachovia Bank, 301 South Tryon St, NC0663, Charlotte, NC 28288, *john.zuo@wachovia.com*

Key Words: Credit Scoring Model, Discriminatory Power, Monte Carlo, KS, Model Validation

The discriminatory power from a credit scoring model refers to its ability to differentiate between good and bad credits. It can be measured by various statistics such as Kolmogorov-Smirnov test, Information Value, Gini Coefficient from a CAP curve, Area under Curve from a ROC curve, and Conditional Information Entropy Ratio. However, the mathematical relationships among those statistics are not available in the literature. This paper demonstrates the empirical relationships among the statistics using Monte Carlo technique. The results can be used as a reference to derive the estimates for other statistics after one test (which may be a favorable test for a practitioner) is calculated during a model validation process.

Anticipating Extreme Changes in the Performance of Long Repayment Term Consumer Loan Portfolios Through Eigenvector Analysis of Markov Chain Matrices

✱Alex Strounine, Strounine Thompson Technology Group, PO BOX 990934, Boston, MA 02199-0934, *alex@strounine.com*

Key Words: consumer credit, Markov chains, recession, credit risk, risk management, forecasting

Predicting credit performance of consumer loans that have long (12 to 20 years) repayment terms can be a challenging task. When loans' past prepayment and default rates are taken as best approximations of the future outcomes, the forecast might be outdated regardless of modeling techniques applied. Competitive activity and changes in the economic environment are common factors that undermine the accuracy of such predictions. We present a technique that uses short term models and Markov chain analysis to monitor inconspicuous movements in pools of new loans to detect trends that signify qualitative changes which will result in extreme deviation of actual loan performance form the predicted one. Whether we expect a recession, an economic crisis, or a revolution, the method provides tools to detect an extreme change sooner and to define strategies to mitigate potential credit loss.

Ruin Probabilities in Risk Model with the General Mixture Distribution

✱Min Deng, Maryville University of St. Louis, 650 Maryville University Drive, St. Louis, MO 63141, *mdeng@maryville.edu*

Key Words: Aggregate Claims, Counting distribution, Claim distribution, Mixture Distribution, Ruin, Surplus

This paper focus on one important aspects of risk theorem—the probability functions of ruin. In this paper we are going to propose a method to deriving recursive and explicit formulas for ruin probabilities in risk model with the general mixture distribution. Some numerical examples to illustrate the results will be given.

Management of Business Processes in Uncertain Conditions

✱Maxim B. Khayrullin, Novosibirsk State Technical University, Shirokay st. 111, flat 16,, Novosibirsk - 136, 630136 Russia, *khairullin@ngs.ru*; Anatoly A. Naumov, Novosibirsk State Technical University

Key Words: Estimating of risks, Factor analysis, Sensetivity analysis, Business processes, Reengineering, Modeling

In the paper the management of economical systems problem with using special models in business process form is investigated. Tasks of choosing more efficient business processes by estimating of risks, sensitivity and factor analysis are considered. Offered methods of business process portfolio optimization with presenting of risks generalize earlier far-famed methods. Conducting of economic systems analysis on sensitivity allows to find the factors which most essentially influence economic parameters. The results of factor analysis a basis to identify so-called bottlenecks in economic system. The complex analysis of economic parameters in view of uncertainty is based on methods of modeling. And on its basis management for problems of maneuvering and reengineering economic systems is developed.

Modeling Dependence in the Design of Crop Insurance Contract: A Semiparametric Copula Model Approach

✱Ying Zhu, North Carolina State University, 2801 Founders Drive, Department of Economics, Raleigh, NC 27695-8110, *yzhu@ncsu.edu*; Sujit Ghosh, North Carolina State University; Barry Goodwin, North Carolina State University

Key Words: copula function, semi-parametric estimation, crop insurance

I evaluate and model multivariate risk factors and their interaction in agricultural production with implications for rating crop insurance contracts. The copula approach is applied to construct a multivariate distribution of the risks of revenue variability that arises from changes in crop prices, yields, or both. Various copula functions are investigated for their suitability in modeling yield and price risks. A semiparametric method is proposed to estimate the joint distribution of price and yield risks by evaluating a parametric copula function at their nonparametric kernel-based marginal distributions. The proposed method avoids the possible misspecification of marginal distributions on the estimation of copula dependence parameters. Estimation results show evidence of efficiency gains over the two-stage MLE of a copula parameter, conditional on the MLE of marginal parameters.

Using Time-Varying GEVs To Model Extreme Financial Returns

✱Mark L. Labovitz, University of Colorado, Denver/Lipper, A Reuters Company, 1205 S. Riverbend Court, Superior, CO 80027, *mark.labovitz@lipper.reuters.com*

Key Words: GEV Distributions, Financial Risk, Financial Portfolio Returns, Modeling Extreme Returns, Equity Extreme Returns, Incorporating Economic Covariates

Precise estimates of risk in portfolio returns are of paramount concern to financial professionals. The need for precision in these estimates is increasing with the increased marketplace volatility presently being observed. General Extreme Value (GEV) distributions are useful in describing the uncertainty of extreme returns. Unfortunately, the common parameterization of these distributions does not take into account changes in the market and economy over time, so extremes under a static parameterization may be considerably more conservative or optimistic than the present state of the market. The author has expanded a model for time-varying parameterization of GEV distributions and has applied it to a worldwide set of publicly traded equities. Using these time-varying parameters, changes may be described as a function of market capitalization, sector, and other economic factors.

Improved Statistical Emphasis on Hedge Fund Trading Strategy and Risk Management

✱Bogong T. Li, Bank of America, N.A., 5321 Strawberry Hill Dr. Apt E, Charlotte, NC 28211, *tim.li@bankofamerica.com*

Key Words: Hedge fund, Risk Model, Portfolio analytics, Rare events, Trading, Pattern recognition

This presentation first provides a general overview of key hedge fund trading strategies and risk models where statistical techniques play a fundamental role. Then it suggests statistical analyses that are necessary to secure profits and to prevent underestimation of extraordinary events that may cause catastrophic failures like those we have witnessed in recent memory. The analyses encompass a broad spectrum of statistical techniques including probability distribution of rare events, statistical pattern recognition, Monte Carlo and data-mining techniques etc. The paper thereby advocates the absolute importance of statistical validation in contrast to intuition, the merit of simple model with clear assumptions verses complex model with mingling assumptions at various stages, and the accuracy as well as the speed of statistical delivery in a dynamic trading environment.

239 Survival Analysis: New Theoretical Developments ▲

Biometrics Section, IMS

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

An Accelerated Failure Time Cure Model for Time-to-Event Data with Masked Cause of Failure

✱Jing J. Zhang, Harvard University, 655 Huntington Avenue, SPH2 4th Floor, Boston, MA 02115, *jjzhang@fas.harvard.edu*; Molin Wang, Dana Farber Cancer Institute and Harvard University

Key Words: accelerated failure time model, cure rate, masked cause, EM algorithm

We consider the analysis of time-to-event data that is subject to a cure rate with masked cause of failure. Assuming an accelerated failure time model with unspecified error distribution for the time to event of interest, we

propose rank-based estimating equations for the model parameters and use a generalization of the EM algorithm for parameter estimation. The motivation comes from an International Breast Cancer Study Group (IBCSG) clinical trial, where we characterize the process of treatment-induced amenorrhea (TIA), complicated by the fact that natural menopause may also occur and is indistinguishable from TIA unless a recovery of menses is observed after treatment end. Moreover, not all patients will experience TIA (cured proportion). A simulation study is conducted to evaluate the performance of the proposed method, and an application to the IBCSG data is undertaken.

Adjust the Estimates of Hazard Ratios for Group Sequential Trials Using the Bayesian Dynamic Survival Model Combined with Computation Techniques

✱Jianghua He, Univeristy of Kansas Medical Center, Department of Biostatistics, Mail Stop 1026, 3901 Rainbow Blvd, Kansas City, KS 66210, *jhe@kumc.edu*; Matthew Mayo, The University of Kansas Medical Center

Key Words: Cox model, non-proprotional hazards, Bayesian Dynamic Survival Model, censoring, group sequential trials

The Cox model is commonly used for survival-analysis-based monitoring of group sequential trials. When the proportional hazards assumption is violated, the large scale of censoring at various time points due to the staggered entry in a trial may strongly affect the estimated hazard ratios. We propose to use the Bayesian Dynamic Survival Model combined with computation techniques to generate hazard ratios adjusted for the effect of staggered entry. The adjusted estimates make full use of information at the interim analysis stage and have better interpretation. An extensive simulation study is applied to compare the performances of the unadjusted and adjusted estimates.

Estimating Cumulative Treatment Effects in the Presence of Nonproportional Hazards and Dependent Censoring

✱Guanghui Wei, The University of Michigan, Department of Biostatistics, U of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109, *ghwei@umich.edu*; Douglas E. Schaubel, The University of Michigan

Key Words: Dependent censoring, IPCW, IPTW, Nonparametric estimtor, Survival analysis, Time-dependent effect

In medical studies featuring survival time data, non-proportional hazards and dependent censoring are very common issues when estimating the treatment effect. We propose a measure for estimating the cumulative treatment effect in the presence of non-proportional hazards. The proposed estimator is computed by first using inverse probability of treatment weighting (IPTW) to balance the treatment-specific covariate distributions. At the next stage, Inverse probability of censoring weighting (IPCW) is applied to handle the dependent censoring due to known prognostic factors. Asymptotic properties of the proposed estimator are derived, and the finite-sample properties are assessed in simulation studies. The proposed method is applied to organ failure data.

Optimal Inferences for Proportional Hazards Model with Parametric Covariate Transformations

✱Chunpeng Fan, sanofi-aventis, Mail Stop: BRW M-203A, 1041 Rt 202-206, P.O.Box 6800, Bridgewater, NJ 08807-0800, *Chunpeng.Fan@sanofi-aventis.com*; Jason P. Fine, University of Wisconsin-Madison; Jong-Hyeon Jeong, University of Pittsburgh

Key Words: Cox proportional hazards model, Parametric covariate transformation, Optimal test, Semi-parametric model, Log-linearity

Traditional Cox model assumes a log-linear relationship between covariates and the underlying hazard function. However, the linearity may be invalid in real data. We propose a generalized Cox model which uses parametric covariate transformations to recover it. While the proposed generalization may seem simple, the inferential issues are challenging due to the loss of identifiability under no effects of transformed covariates. Optimal tests are derived for certain alternatives. Rigorous parameter inference is established under regularity conditions and non-zero transformed covariate effects. The estimates perform well in simulation studies with realistic sample size and the proposed tests are more powerful than the usual or sup partial likelihood ratio test. A real data is used to illustrate the model building and the better fit of the proposed model, comparing to traditional Cox model.

Discriminating Weibull and Log-Normal Under Type-II Censored Data

✱Arabin K. Dey, IIT Kanpur, Department of Mathematics and Statistics, IIT Kanpur, Kanpur, International 208016 India, *arabin@iitk.ac.in*; Debasis Kundu, IIT Kanpur

Key Words: type-II censoring, asymptotic distributions, likelihood ratio test, probability of correct selection

In this paper we consider the problem of discriminating between Weibull and log-normal distributions when the data are Type-II censored. We use the ratio of maximized likelihoods in choosing between the log-normal and Weibull distributions. We obtain the asymptotic distributions of the logarithm of the ratio of maximized likelihoods. It is used to determine the probability of correctly selecting the parent distribution. Asymptotic results are verified in case of small sample size with some simulated results. Two data analysis are also performed for illustrative purpose.

Backward Estimation of Medical Cost in the Presence of a Failure Event

✱Kwun Chuen (Gary) Chan, Johns Hopkins University, 615 North Wolfe Street, Department of Biostatistics, Baltimore, MD 21205, *kcchan@jhsph.edu*; Wang Mei-Cheng, Johns Hopkins University

Key Words: Marked Process, Left Truncation, Prevalent Cohort, Right Censoring, Stochastic Process

A backward medical cost process is introduced for modeling end-of-life-cost. Estimator for mean backward cost function is proposed for right censored and possibly left truncated data. The proposed estimator converges weakly to a Gaussian process with an easily estimated covariance function. The novelty of this approach is to consider the partially-observed failure event to be the time origin for the medical cost process. Besides medical cost, backward process can be defined for any measurements and the proposed methods can also be applied to those cases. We use medical cost as an example throughout this article and illustrate the proposed methodologies by analyzing the SEER-Medicare linked data. The results show that among ovarian cancer patients, half of the final year of life cost is spent in the final three months of life.

A New Partial Likelihood for the Cox Proportional Hazards Model

✱Liyuan (Larry) Ma, Merck & Co., Inc., West Point, PA 19486, *larry_ma@yahoo.com*; Devan V. Mehrotra, Merck Research Laboratories; Woollcott Smith, Temple University

Key Words: Partial Likelihood, Cox Proportional Hazards Model, Proportional Odds, Conditional Probability, Finite Sample, Two-Sample Problem

The Cox (1972) proportional hazards model is based on the concept of the partial likelihood (Cox 1975), and this ingenious approach leads to the elimination of the infinite dimensional baseline hazard function from the estimation of regression parameters with censored data. We show that under the proportional odds assumption, Cox partial likelihood is indeed a product of conditional probabilities, which are free of nuisance parameters. However, under the proportional hazards assumption, the nuisance parameters cannot be removed from the partial likelihood. We develop a new partial likelihood for the proportional hazards model and show that for finite samples, the efficiency of the parameter estimates under proportional hazards can be greatly improved. Asymptotically, the new partial likelihood converges to the Cox partial likelihood.

240 Statistical Issues in Oncology Trials ●▲

Biopharmaceutical Section, Biometrics Section
Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Family of Optimum M-Stage Designs for Single Arm Phase II Oncology Clinical Trials

✱Muhammad Jalaluddin, Novartis Pharmaceuticals, 180 Park Avenue, Bldg 104/2J24, Florham Park, NJ 07932, *muhammad.jalaluddin@novartis.com*

Key Words: Optimum design, clinical trials, multi-stage design, two-stage design

Optimum m-stage designs for single arm studies are commonly used in oncology phase II programs. Existing methods select designs with minimum expected sample size under null hypothesis (E(n|H0)), which is denoted by minN. These approaches miss designs with better properties when corresponding E(n|H0) is slightly larger than minN. This issue has been addressed by introducing a concept of family of optimum designs. For a given hypothesis testing setup and a non-negative real number D, a family of optimum m-stage designs consists of all m-stage designs with E(n|H0) less than or equal to minN + D. One can chose a design from this family based on his/her selection criteria. For example, a design with maximum probability of early termination when H0 is true, smallest sample size in stage 1, minimax, etc. Methodologies associated with this approach along with examples will be presented.

Relationship Between Type I Error and Frequency of Progression Assessment

✱Xiaoping Jiang, U.S. Food and Drug Administration, 10903 New Hapmshire Ave, SIlver Spring, MD 20993, *xiaoping.jiang@fda.hhs.gov*; Kun He, U.S. Food and Drug Administration; Rajeshwari Sridhara, U.S. Food and Drug Administration

Key Words: Assessment of Disease progression, type I error, Power, simulations, cancer, disease growth rate

Time to progression defined as the time from randomization to disease progression, is commonly used as an endpoint in evaluating drug products to treat cancer. Progression is measured using pre-defined criteria at preplanned frequency of visits such as every 6 weeks or every 12 weeks. Progression assessments are expensive and therefore disease monitoring frequency is generally guided by clinical practice. The natural history of the disease may be either aggressive or slow growing tumors. In this study we examine the relationship between type I error and different frequencies of assessment in aggressive disease versus slow growing disease using simulated data assuming no censored events and varying percentage of censored events.

Consequences of Asymmetry in Progression Assessments

✱Shenghui Tang, USFDA, 10903 New Hampshire Avenue, WO #22, Rm 1209, USFDA, Silver Spring, MD 20993, *shenghui.tang@fda.hhs.gov*; Chris Holland, USFDA; Rajeshwari Sridhara, U.S. Food and Drug Administration

Key Words: Asymmetry, Simulation, False Positive Rate, Disease Growth Rate, Cancer

Progression-free survival defined as the time from randomization to either progression or death, is one of the efficacy endpoints used in evaluating oncology drug products. Disease progression is generally measured using predefined criteria at preplanned frequency of visits. Some studies have shown that asymmetry in progression assessments between the treatment arms can introduce bias and confound the true treatment differences. In this study we examine the impact of asymmetry with respect to false positive rate using simulations when (1) the assessments are more frequent versus less frequent, (2) the disease under consideration is fast growing versus slow growing disease, and (3) percentage of censored data varies.

Evaluation of Time-to-Progression: Where Your Censor Matters?

Chia-wen Ko, U.S. Food and Drug Administration; ✱Somesh Chattopadhyay, U.S. Food and Drug Administration, 10903 New Hampshire Ave, Bldg 22, Rm 1211, Silver Spring, MD 20993, *somesh.chattopadhyay@fda.hhs.gov*; Rajeshwari Sridhara, U.S. Food and Drug Administration

Key Words: disease progression, missing data, censoring, simulation, type I error, cancer

Time to progression (TTP), defined as time from randomization to disease progression (PD), is an important endpoint in evaluating efficacy of cancer drug products. PD is measured at pre-planned patient visits and not on a daily basis. Missed PD assessments (MA) occur for reasons such as unevaluable measurements or missed visits. When PD is documented immediately after having MA, there are options in determining TTP: censor observation at the date of last adequate assessment visit with no PD, count observation as an event on the date of documented PD, consider as an interval-censored observation, and others. We will examine these options for the effect on type I error rate, power, and bias in estimation using simulated data with various MA patterns between treatment arms. Advantages and disadvantages of each of the options will be discussed.

Planning for Phase III Oncology Trials Without Adequate Data from Phase II Trials

✱Pralay Mukhopadhyay, Bristol-Myers Squibb Company, 5 Research Pkwy, P.O. Box 5100, Wallingford, CT 06460, *pralay.mukhopadhyay@bms.com*; Thomas Kelleher, Bristol-Myers Squibb Company

Key Words: PFS, OS, futility, PET

Phase III oncology trials are traditionally designed with progression free survival (PFS) or overall survival (OS) as the primary endpoint. These trials are usually very large and are often conducted without adequate efficacy data in the phase II setting. Commonly a different endpoint is used in a phase II setting (such as tumor response) which may not help in providing assumptions for a PFS or OS benefit. In these situations strict futility rules at earlier stages of the phase III trial may be considered. We investigate sample sizes that will be required in a randomized phase II setting (with similar endpoints as phase III) to adequately rule out lack of benefit. Using these criteria as futility rules in the phase III setting we investigate the loss in power and probability of early termination (PET) compared to running a separate phase II study through simulation.

Dose Reduction Profiles for Combination Drug Studies

✱John J. Peterson, GlaxoSmithKline Pharmaceuticals, Research Statistics Unit (UP-4315), 1250 So. Collegeville Road, Collegeville, PA 19426, *john.peterson@gsk.com*

Key Words: antagonism, graphics, Minto-White model, oncology, response surface methodology, Synergy

In combination drug studies, particularly in oncology, it is important to find a combination of drugs that provide good efficacy but with reduced toxicity. This can sometimes be achieved if each drug has a different primary side-effect. This paper proposes "dose reduction profiles" which show what dose combinations are associated with the best dose reductions for each drug in combination relative to the amount of each drug alone needed to obtain the same level of efficacy. A connection with Loewe synergy is also established.

Sensitivity Analysis for Treatment Drop-In in Oncology Clinical Trials

✱Cheng A. Rong, Amgen, Inc, One Amgen Center Drive, Thousand Oaks, CA 91320, *crong@amgen.com*; Michael Wolf, Amgen, Inc; Zhiying J. Pan, Amgen, Inc; Xiang C. Zhang, Amgen, Inc

Key Words: Treatment Drop-in, Treatment Crossover, Overall Survival, Oncology Clinical Trials

In early-line oncology clinical trials, evaluation of treatment effect of current line therapy on overall survival (OS) is often confounded by the issue of treatment drop-in, where patient enter into next line of treatment after they fail the current treatment due to disease progression. In this presentation, we will discuss several statistical methods on how to assess current line treatment effect on OS adjust for treatment drop-in as well as on how to address the potential impact of treatment drop-in to the comparison of OS. The sensitivity analysis will focus on several issues: first, how to set up reasonable assumptions for treatment drop-in; Secondly, how to assess whether the current treatment is truly effective; Thirdly, if it does, what are the potential ranges of treatment effect for the current therapy after adjusting for the treatment drop-in. Examples will be provided.

241 Mining Complex and High-Dimensional Data

Section on Statistical Computing, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, Section on Statisticians in Defense and National Security, IMS

Tuesday, August 5, 8:30 a.m.–10:30 a.m.

A Comparison of Several Measures of the Center of a Functional Data Set

✱David B. Hitchcock, University of South Carolina, Department of Statistics, 216 LeConte College, Columbia, SC 29208, *hitchcock@stat.sc.edu*

Key Words: Median, Curves, Distance, Data depth, Central tendency

Among several relatively intuitive choices, we compare possible measures of the center of a functional data set (i.e., a set of observed curves). We discuss advantages and disadvantages of estimating the "central curve" underlying a data set with: (1) the pointwise mean curve; (2) the pointwise median curve; (3) a pointwise trimmed mean curve; and (4) the "most central" sampled curve. The goodness of these curve estimators is judged based on several characteristics, including distance to the true center, total distance to all sample curves, and how well the curvature of the measure of center matches that of the true central curve and that of all sample curves generally. We present results of a simulation study to evaluate the four competing measures. A suggestion is offered for relating these results to a possible functional-data version of Hogg's 1967 robust estimation method.

Statistical Inference with High-Dimensional Data

✱Min Hee Kim, The Pennsylvania State University, 219 S. Sparks St. Apt 6, State College, PA 16801, *mzk132@psu.edu*; Michael G. Akritas, The Pennsylvania State University

Key Words: Hard thresholding, thresholding parameter, ordered thresholding

Hard thresholding is a method for increasing the power of test procedures in settings where the number of parameters being tested increases with the sample size. We will motivate the need for improved power in the simple case of testing a simple null versus simple alternative. Then, we will consider a simple null versus composite alternative and introduce hard thresholding. The important part of this procedure is to choose the thresholding parameter. With a large thresholding parameter, hard thresholding procedure does not work well. In this paper, we will propose the test statistic based on the linear combination of order statistics, called the ordered thresholding statistic. The asymptotic theory works better for the ordered thresholding. The order thresholding also has the highest power, when we compare the efficiency of testing procedures, via simulation.

Tracking the Best Predictor with a Detection-Based Algorithm

✱Yannig Goude, EDF R&D, 1 avenue du général de Gaulle, Clamart, International 92140 France, *yannig.goude@edf.fr*

Key Words: online learning, break detection, combining forecasts, time series, oracle inequalities, aggregating

We consider the problem of online learning a sequence of stochastic observations for which we can make a finite number M of predictions with a fixed set of experts. The aim is to combine them in order to achieve, under general assumptions, the predictive performance of an a-priori unknown referee called the Oracle predictor, classically the best predictor among M. We study the case of breaks. We suppose that the best predictor can change with time and that the sequence of the optimal predictor is the Oracle. Machine learning algorithms exist that achieve this Oracle performance in a non-stochastic context. We adopt a stochastic point of view and propose a new family of algorithms based on break detection methods. We obtain upper bounds on the loss and risk of the corresponding mixing predictors. We also propose different simulation studies that show the interest of these new algorithms.

Classification of Alzheimer's and Normal fMRI Scans Using Temporal Network Distance Matrices

✱Ariana Anderson, University of California, Los Angeles, 8125 Math Sciences Bldg., Box 951554, Los Angeles, CA 90095-1554, *ariana@stat.ucla.edu*; Ivo D. Dinov, University of California, Los Angeles

Key Words: fMRI, medical imaging, classification, machine learning, data mining, pattern recognition

fMRI is an imaging method that records blood flow patterns over space and time. Analysis is impeded by having a weak signal embedded in a high-dimensional space, where a single scan can contain over 15 million recordings. We will outline a method for classification and discrimination among distance matrices, where each matrix is representative of the divergence of spatial network signals fluctuating over time within a single Alzheimer's or normal subject. A computational framework for classification of objects based on distance matrices will be proposed, extending to cases where the matrices are of different dimensions and represent unique underlying manifolds. Geometric properties of each distance matrix will be used to discriminate between groups, to create a cohesive method for classification of objects based on distance matrices representing disparate manifold.

Linear Dimension Reduction for Multiple Multivariate Skew-Normal Densities

✱Philip D. Young, Baylor University, One Bear Place, Department of Statistical Science, Waco, TX 76798-7140, *Philip_Young@baylor.edu*; Jack Tubbs, Baylor University; Dean Young, Baylor University

Key Words: Linear feature selection, Singular value decomposition, Expected probability of misclasssification

We derive a new linear dimension-reduction method to determine a low-dimensional hyperplane that preserves or nearly preserves the original feature-space separation of the individual populations and the Bayes probability of misclassification for multiple multivariate skew-normal populations. Our new dimension-reduction method gives a linear feature-selection transformation for low-dimensional representation. Also, we give necessary and sufficient conditions which determine the smallest reduced dimension that retains the Bayes probability of misclassification from the original full-dimensional space in the reduced or transformed space. Finally, we demonstrate the application of our new procedure with several examples.

Principal Component Rotation in High Dimensions

✱Trevor Park, University of Florida, Department of Statistics, 102 Griffin-Floyd Hall, P.O. Box 118545, Gainesville, FL 32611, *tpark@stat.ufl.edu*

Key Words: penalized likelihood, rotation

Principal component estimates are subject to high sampling variability, especially when dimension exceeds sample size. This variability can be reduced by using methods for simplifying the components, including methods based on dimension reduction, thresholding, and penalization. Some recent general methodology will be reviewed in the high-dimensional context, and a new penalty-based rotation method will be proposed, as an extension of current methods. Some computational details will be described.

Small Sample Inference for Generalization Error in Classification

✱Eric B. Laber, The University of Michigan, 1085 South University, Department of Statistics, Ann Arbor, MI 48104, *laber@umich.edu*; Susan A. Murphy, The University of Michigan

Key Words: classification, training bounds, small sample, machine learning

Confidence measures for the generalization error are crucial when small training samples are used to construct classifiers. A common approach is to estimate the generalization error by resampling and then assume the resampled estimator follows a known distribution to form a confidence set. Alternatively, one might bootstrap the resampled estimator of the generalization error to form a confidence set. In contrast, we construct a confidence set for the generalization error by use of an upper bound on the deviation between the resampled estimate and generalization error. The confidence set is formed by bootstrapping this upper bound. In cases in which the approximation class for the classifier can be represented as a parametric additive model, we provide a computationally efficient algorithm. This method exhibits its superior performance across a series of test and simulated data sets.

242 Contributed Poster Presentations

Biometrics Section, Biopharmaceutical Section, ENAR, Section on Health Policy Statistics, Section on Statistical Computing, Section on Statistics in Epidemiology, WNAR

Tuesday, August 5, 8:30 a.m.–10:20 a.m.

Bayesian Array Analysis To Assess the Effects of Nutrient Deficiencies in White Lupin

✱Robert Norton, California State University, East Bay, 25800 Carlos Bee Blvd, Hayward, CA 94542, *rnorton@horizon.csueastbay.edu*; Mousumi Rath, California State University, East Bay; Bandita Parhy, California State University, East Bay; Claudia Stone, California State University, East Bay

Key Words: RNAi, array analysis, biometrics, bioinformatics, computational biology, qPCR

Lupinus albus L. has unique adaptations for extracting phosphorus (P) and iron (Fe) from soil when other plants cannot. The objective of this research is to identify genes up- or down-regulated in proteoid roots under P, Fe and nitrogen (N) deficiency. We begin with array data for two up-regulated genes, six replicates exist for each gene under each condition. Assume that the number of ratios of treatment to control expression intensities are>2; is a binomial variable with unknown proportion pi for each gene. This paper compares outcomes using a standard chi-square test; a flat prior on both unknown proportions; and a beta prior on pi. Since the number of replicates is small, it is reasonable to expect Bayesian results to be more useful in making decisions about the outcomes. The Bayesian analysis provides estimates for the proportion pi and the covariance of the two gene expressions.

Spatial Event Cluster Detection Using a Normal Approximation

✱Rhonda Rosychuk, University of Alberta, 4th Floor, Rm 9423 Aberhart Centre, 11402 University Avenue NW, Edmonton, AB T6G 2J3 Canada, *rhonda.rosychuk@ualberta.ca*; Mahmoud Torabi, University of Alberta

Key Words: disease cluster detection, normal approximation, surveillance

Tests for the detection of geographic clusters seek to identify regions that have higher disease rates than expected. These methods are generally applied to cases of disease, but surveillance of disease-related events may also be of interest. Recently, a compound Poisson approach that detects event clusters by testing individual areas (that may be combined with neighbors) was proposed. However, the required probabilities are obtained from a recursion relation that can be cumbersome if the number of events is large or analyses by strata are performed. We propose a simpler approach that uses a normal approximation. This method is easy to implement and is applicable to situations where the population sizes are large and the population distribution by important strata may differ by area. We illustrate the approach on pediatric self-inflicted injury presentations to emergency departments.

Statistical Methods for Automated Drug Susceptibility Testing: Minimum Inhibitory Concentration Prediction from Growth Curves

✱Xi (Kathy) Zhou, Cornell University, 411 East 69th Street, New york, NY 10021, *kaz2004@med.cornell.edu*; Merlise Clyde, Duke University; James Garrett, Becton Dickinson Diagnostic Systems; Viridiana Lourdes, Morgan Stanley; Michael O'Connell, Insightful Corporation; Giovanni Parmigiani, Johns Hopkins University; David J. Turner, Becton Dickinson Diagnostic Systems; Timothy M. Wiles, Becton Dickinson Diagnostic Systems

Key Words: decision theory, BIC, model selection, Bayes, MIC estimation

Rapid drug resistance characterization of bacterial strains through minimum inhibitory concentration (MIC) determination is crucial for managing patients with infections. Here we present a novel probabilistic approach for accurate MIC prediction using features of bacterial growth curves. Bayesian model selection was used to select the most important features that predict bacterial growth. A probabilistic model was developed to estimate the probability of a given drug dilution being the MIC based on growth. Decision theoretical MIC estimates can be obtained by adjusting penalty weights for various errors that reflect FDA regulatory requirements. Our methods are easily automated and have been incorporated into the Becton-Dickinson PHOENIX automated susceptibility testing system that rapidly and accurately classifies the resistance of a large number of micro-organisms in clinical samples.

Comparisons of Titer Estimation Methods for Multiplexed Pneumococcal Opsonophagocytic Killing Assay

✱Deli Wang, The University of Alabama at Birmingham, *deli.wang@ccc.uab.edu*

Key Words: Multiplexed opsonization assay, Four parameters logistic model, Titer estimation

Titer estimation is one of the major components of immunoassay and vaccine development. A multiplexed In vitro opsonization assay (MOPA) is widely accepted to quantitate Streptococcus pneumococcal antibodies to serotype-specific pneumococcal capsular polysaccharide. Titer estimation of OPA is one important component of standardization of OPA and the selected statistical method is a factor influencing accuracy and precision of titer estimation. We evaluated five titer estimation methods for pneumococcal OPA in terms of precision and accuracy using three data sets generated by specifically designed experiments. We concluded that the traditional direct method did not perform as well as the other four methods in terms of precision and accuracy of titer estimation. The 4PL method is an alternative choice for OPA titer estimation. UAB opsotiter was developed to implement OPA titer estimation.

New Bimodel Classifier for Predicting Outcomes of Prostate Cancer Patients

✱Zhenyu Jia, University of California, Irvine, 36 Oxford, Irvine, CA 92612, *zjia@uci.edu*; Yipeng Wang, The Sidney Kimmel Cancer Center; James Koziol, The Scripps Research Institute; Michael McClelland, The Sidney Kimmel Cancer Center ; Dan Mercola, University of California, Irvine

Key Words: ANOVA, classifier, cross validation, prostate cancer

The majority of prostate cancer cases are "indolent" that do not threaten lives. In order to improve disease management, reliable molecular indicators are needed to distinguish the indolent cancer from the cancer that will progress. Conventional methods can not apply to this study because the tissue samples are very heterogeneous in cell composition. Here we illustrate the expression level of any gene by a linear model considering the contributions from four principal types of cells and their interactions with aggression indicators (known relapse or nonrelapse status of the cases). ANOVA is used to identify cell specific relapse associated genes that possess discriminative power. The expression patterns of those selected genes may be described using two Gaussian models, one for relapse and one for nonrelapse cases. Thus they can be used for predicting outcomes of newly diagnosed cases.

Locating DNA Copy Number Changes by Using a Statistical Change Point Model

✱Paul Plummer, University of Missouri-Kansas City, Department of Mathematics and Statistics, 206 Haag Hall, UMKC, 5100 Rockhill Rd, Kansas City, MO 64110-2499, *pplummer@ucmo.edu*

Key Words: DNA copy number, change point, hypothesis testing, asymptotic distribution, aCGH profiles

The importance of DNA copy number changes in the study of cancer cell development is well documented. In this study the statistical mean change point model with variance known is used to determine the location of copy number changes in microarray comparative genomic hybridization (aCGH) profiles. The p-value is calculated using the asymptotic null distribution of a likelihood ratio based test statistic. The approach is applied to fifteen fibroblast cancer cell lines and two breast tumor cell lines. The application results indicate that the proposed approach is effective for identifying copy number changes (break points). The drawbacks of applying a single change point model with binary segmentation to a data set with multiple change points are also discussed.

Estimating the Proportion of Differentially Expressed Genes in Comparative DNA Microarray Experiments

✱Ching-Ray Yu, Pfizer, Inc., 685 East 44th ST, 685/13/06, New York, NY 10017, *ching-ray.yu@pfizer.com*

Key Words: biological process, DNA microarray, differentially expressed genes, toxicology experiment

DNA microarray experiments, a well-established experimental technique, aim at understanding the function of genes in some biological processes. One of the most common experiments in functional genomics research is to compare two groups of microarray data to determine which genes are differentially expressed. In this paper, we propose a methodology to estimate the proportion of differentially expressed genes in such experiments. We study the performance of our method in a simulation study where we compare it to other standard methods. Finally we compare the methods in real data from two toxicology experiments with mice.

Discovery of Novel Protein Domains: The Parasitic Connection

✱Ian Lee, National University Hospital, Department of Haematology-Oncology, 5 Lower Kent Ridge Road, Singapore, International 119074 Singapore, *Ian_HH_LEE@nuh.com.sg*

Key Words: HMM, Gibbs Sampler, Convergence, Parasites

Domains are key indicators of protein function. While statistical models based on iterated profile searches (PSIBLAST) have been instrumental in their detection, they face certain limitations. Convergence is hampered by inclusion of closely-related sequences that limit the scope of the model or insufficient signal strength to detect distant relationships. Recent large-scale sequencing of parasite genomes provides a means of alleviating this problem. Using sequences from the malaria and trypanosome genomes as bridging points in model construction, we have previously identified the RAP and SMP domains (Lee and Hong 2004, 2006). We further illustrate this with identification of a novel domain through the use of another sequenced parasite genome: Entamoeba histolytica. The domain is amplified in this species and other microbes and is of potential relevance for human pathology.

Identifying Allele-Specific Gene Expression in Human Brains Using the Illumina Sentrix Array Matrix and the Goldengate Assay

✱Xin V. Wang, University of California, Berkeley, 367 Evans Hall, Berkeley, CA 94720-3860, *xvwang@stat.berkeley.edu*; Terence P. Speed, University of California, Berkeley; Charles E. Glatt, Cornell Medical School

Key Words: Illumina, allele-specific gene expression, empirical Bayes

Allele-specific gene expression (ASE) occurs when expression levels of the two alleles differ. This phenomenon has been shown to be relatively common and may be important in gene regulation. The Illumina platform is used to identify such allelic imbalance in three brain regions from 85 human brains, 11 of which are duplicated. Approximately 1000 genes represented by 1510 SNPs are queried. We will present the methods and results used in the analysis of this data set. In particular, background correction and normalization steps that are specific to this platform are taken before downstream analysis. An empirical Bayes approach is then used to identify genes that exhibit ASE in a subset of the human brains. This approach is shown to be more reliable than a method previously used for analyzing this type of data when assessed by the concordance of SNP pairs from the same coding regions.

A Genetical Genomics Approach to Genome Scans for Complex Traits

✱Guoying Sun, The University of Georgia, Athens, GA 30602, *gysun@stat.uga.edu*; Paul Schliekelman, The University of Georgia

Key Words: eQTL, EM Algorithm, causative loci

We have developed a EM-algorithm based method for testing whether the controlling factor for a group of highly correlated genes is also a causative locus for a disease or other clinical trait. We use this as part of a new strategy for genome scans for complex traits. We first screen the data to identify the set of expression levels which are the most promising targets for association with the trait. We next use hierarchical clustering to identify groups of highly correlated transcripts. We apply our EM algorithm method to each gene group to test whether any controlling loci are also causative for the clinical trait. Expression QTLs for such gene groups become candidate loci for the clinical trait. Using both simulated and real data, we show this strategy to have excellent power relative to existing methods.

Rediscovering the Power of Well-Planned Contrasts: Normalization and Analysis of cDNA Microarray Using Linear Contrasts

✱Liping Huang, University of Kentucky, Departmen of Statistics, 815 Patterson Office Tower, Lexington, KY 40508, *liping@ms.uky.edu*; Naoki Miura, University of Kentucky; Michael Mienaltowski, University of Kentucky; James MacLeod, University of Kentucky; Arnold Stromberg, University of Kentucky; Arne Bathke, University of Kentucky; Constance Wood, University of Kentucky

Key Words: cDNA microarray, dye bias, linear contrast

In cDNA microarray experiments, there are many sources of systematic variation. Removal or modeling of these variations is important for comparing the gene expression level across experiments. We adapted a planned linear contrasts approach to remove dye bias and between-chip effects and then constructed a sequence of contrasts to identify probe sets with differential expression level. We exemplified our methods in two datasets. The first dataset is a 2x2 factorial design, and the second one is a comparison between two tissues. When used prudentially, the linear contrast approach constitutes a powerful and transparent alternative to a mixed model analysis.

Large-Scale Comparison of Methods for Analysis of Microarrays: Ranked List Accuracy and Sample Size

Robert H. Podolsky, Medical College of Georgia; ✱Nikhil Garge, Medical College of Georgia, Center for Biotechnology and Genomic Medicine, 1120 15th St., CA-4100, Augusta, GA 30912, *ngarge@mail.mcg.edu*

Key Words: microarrays, Optimal Discovery Procedure, shrinkage, list accuracy

The analysis of microarrays involves two related steps: ranking genes in order of the likelihood of a difference between conditions; and choosing some threshold for identifying the genes of interest. We focus on the first step here by conducting a large scale simulation study based on observed effect sizes in three microarray studies. Several methods are compared for their ability to rank genes with regard to being differentially expressed. We also examine the relationship between sample size and accuracy of the ranked list of genes. Our results indicate that while the Optimal Discovery Procedure (ODP) does rank genes more accurately in some cases, shrinkage methods are more consistent across a range of conditions. Current sample sizes do provide sufficient accuracy for ranking the top 10-50 genes, but suffer in accurately ranking the remaining genes.

A Method for Learning Gene Association Networks from High-Dimensional Data

Jie Cheng, GlaxoSmithKline; Xiwu Lin, GlaxoSmithKline R&D; ✱Kwan Lee, GlaxoSmithKline, 19426, *kwan.lee@gsk.com*

Key Words: partial correlation matrix estimation, gene association network, graphical Gaussian models, Bayesian networks

Inferring large scale gene networks from limited continuous data is a challenging problem in bioinformatics. This problem is closely related to partial correlation matrix estimation and graphical Gaussian model learning. We developed a java based tool for such task based on our previous work on learning Bayesian network from multinomial data. The basic idea is to learn the network structure in three phases: drafting, thickening and thinning. In drafting phase, marginal correlations are used to generate an initial guess of the structure. In thickening and thinning phases, low order partial correlations are used to constantly modify the network in a parsimony manner (in terms of both the number of tests and the order of the tests), guided by the network structure at that moment. Simulation data and public functional genomics data are used to evaluate the performance of the proposed method.

Statistical Analysis for Phylochip Data: Acidobacteria Division from Soil

✱Ozgur Turkmen, Auburn University, Department of Mathematics and Statistics, 221 Parker Hall, Auburn, AL 36849, *turkmoz@auburn.edu*; Mark Liles, Auburn University; Brian Manske, University of Wisconsin; Robert Goodman, University of Wisconsin; Nedret Billor, Auburn University

Key Words: Phylochip, Multiple Comparision

A microbial community census involves the laborious and costly cloning and sequencing of the 16S ribosomal RNA gene. A microarray (“phylochip”) with thousands of DNA probes targeting distinct rRNA regions can generate a rapid microbial community profile. A phylochip was developed targeting diverse lineages of soil Acidobacteria (which are abundant in soils yet not cultured) with over 45,000 unique probes. Cloned rRNA genes were used as positive controls. Adding 2.5M betaine to hybridization buffer reduced background signal (2 group comparison tests). All 5 control rRNA genes were correctly predicted based on phylochip data (pairwise differences by multiple testing). Phylochips were used to assess Acidobacteria diversity in a Wisconsin soil, revealing two dominant subgroups. A soil clay fraction had an altered Acidobacteria community, which was confirmed with another statistical method.

Mining Illumina Chip Time Course Microarray Data: Assessment of Normalization Methods and Pattern-Based Clustering of Genes Using Contrast Analysis

✱Suman Duvvuru, University of Tenneessee, 3700 Sutherland Avenue Apt k2, Knoxville, TN 37919, *suman.duvvuru@gmail.com*

Key Words: normalization, clustering, contrast

Statistical genomics is subject to challenges of microarray data analysis including appropriateness of normalization and assessment of statistical significance in a context of massive correlation among statistical tests. Here we present a toolchain for the analysis of time course microarray data which involves two steps: i) Preprocessing of the data: Selection of best normalization method and ii) clustering based on polynomial and Helmert contrast analysis. Several normalization methods such as background correction, average normalization, rank invariant, quantile and cubic spline normalization have been compared and assessed using intraclass correlation. Intraclass correlation is the amount of variance among the different time points relative to total variance. Once the approppriate normalization method has been chosen,contrast analysis is performed to explore the patterns across time.

Modified Linear Discriminant Analysis Approaches for Classification of High-Dimensional Microarray Data

✱Ping Xu, University of Louisville, 555 S. Floyd St., Louisville, KY 40202, *p0xu0001@gwise.louisville.edu*; Guy Brock, University of Louisville; Rudolph Parrish, University of Louisville

Key Words: microarry data, classification, lineat discriminant analysis

Linear discriminant analysis (LDA) is one of the most popular methods of classification. In high-dimensional microarray data classification, due to the small number of samples and large number of features, a problem with classical LDA is the singularity and instability of the covariance matrix. we applied two modified LDA approaches to microarray classification and compared their performance with other popular classification algorithms using both simulated and real datasets. our results show that the overall performance of the two modified LDA approaches is as competitive as support vector machines and better than diagonal linear discriminant analysis, k-nearest neighbor, and classical LDA. It was concluded that the modified LDA approaches can be used as effective classification tools in limited sample size and high-dimensional microarray classification problems.

Repeated Measured Methodology for Spatial Cluster Detection While Accounting for Moving Locations

✱Andrea J. Cook, Group Health Center for Health Studies, 1730 Minor Avenue, Suite 1600, Seattle, WA 98101, *cook.aj@ghc.org*; Yi Li, Harvard University/Dana Farber Cancer Institute

Key Words: Asthma, Cumulative Residuals, Repeated Measures, Spatial Cluster Detection, Wheeze

Spatial cluster detection has become an important methodology in quantifying the effect of hazardous exposures. Previous methods have focused on cross-sectional outcomes that are binary or continuous. There are virtually no spatial cluster detection methods proposed for longitudinal outcomes. This manuscript provides an extension of the cumulative geographic residual method proposed for censored outcomes to repeated measured outcomes. A major advantage of this method is its ability to readily incorporate information on study participants relocation, which most cluster detection statistics cannot. Application of these methods will be illustrated by the Home Allergens and Asthma prospective cohort study analyzing the relationship between environmental exposures and repeated measured outcome, occurrence of wheeze in the last 6 months, while taking into account mobile locations.

Statistics in Evaluating Trauma Center (TC) Performance

✱Wei Xiong, University of Toronto, Queen Wing, 3N-082, 30 Bond Street, Toronto, ON M5B 1W8 Canada, *xiongw@smh.toronto.on.ca*; Khumar Huseynova, University of Toronto; Sandra Goble, NTDB, American College of Surgeons; Avery Nathens, University of Toronto

Key Words: Statistics (Z, W, O/E), Confidence Interval, Zero Observed Deaths, Performance Evaluation, Kappa

For the purposes of external benchmarking of TC care, it is important to identify high/low performing TCs by comparing the ratio of observed versus expected deaths (O/E). A TC with CI (O/E) above/below the value of one is a low/high performing center. The most frequently used methodology in literature uses a Z (TRISS) or W (excess number of survivors /100 patients) statistic. This study compares an O/E statistic with CI based on Chi-square and Poisson distribution. Unlike the usual CIs for O/E in the literature, the proposed method correctly produces reasonable CI for zero observed events. To validate this methodology, we evaluated the concordance of this O/E statistic in a study of 97 TCs across the US. The concordance between O/E and Z (or W) is excellent with Kappa coefficient from 0.81 to 1.00. This study provides a methodology for benchmarking TC with zero observed events.

Evaluating the Hawaii Demonstration To Maintain Independence and Employment

✱Tammy Tom, University of Hawaii, 1015 Wilder Ave #305, Honolulu, HI 96822, *tammytom@hawaii.edu*

Key Words: social security, disability prevention, diabetes, healthy workforce, experimental design, health status

Many workers with potentially disabling health conditions reach a point at which they can no longer work and must rely on federal disability benefits. If workers had access to a comprehensive package of health care services and employment supports, could they postpone, or avoid the need for, disability benefits? The federal Demonstration to Maintain Independence and Employment (DMIE) was developed to study this question and requires funded states to conduct independent evaluations of their projects. The Hawaii DMIE Evaluation uses an experimental design to study the effect diabetes support services have on work-related and health variables.

An Investigation of the Relationships Between Intervention Bills and Child Obesity Rate

✱Melanie Logue, The University of Arizona, AZ , *asumain@yahoo.com.hk*; Margaret G. Ericksen, Arizona State University; Chong Ho Yu, Arizona State University

Key Words: obesity, intervention bills, regression, Parallel coordinate, legal, health policy

The prevalence of childhood obesity has increased dramatically in the past 20 years. The magnitude of childhood obesity has prompted discussions on policy and legal changes as mechanisms for addressing this issue. This objective of this project is to study how introduction of intervention bills affects the child obesity rate (COR) in all U.S. states. Based on the data set released by Health Resources and Services Administration (2005), it was found that the percentage of passed intervention bills is positively associated with COR; however, a negative association was found among the states that passed a moderate percentage of intervention bills. Finding long-term solutions to childhood obesity may require increased dedication and a concerted effort to utilize programs, policies, and laws that will address the multi-faceted nature of the issue.

Fitting a Semi-Markov Process Model to Data on Transitions Between Health States in the Presence of Left Censoring

✱Nathaniel Schenker, National Center for Health Statistics, National Center for Health Statistics, 3311 Toledo Rd, Rm 3209, Hyattsville, MD 20782, *nschenker@cdc.gov*; Liming Cai, National Center for Health Statistics; James Lubitz, National Center for Health Statistics (retired)

Key Words: Aging, Disability, Life Expectancy, Missing Data, Multi-State Life Table, Stochastic EM Algorithm

This presentation describes a method for fitting a semi-Markov Process model to data on transitions between health states (e.g., good health, disability, death) in the older population, when the data are left-censored in the sense that the time of entrance into the initial state observed for each person in the study is unknown. An analog to the stochastic EM algorithm is used to address this missing-data problem. The approach is applied to data from the Medicare Current Beneficiary Survey and the Cardiovascular Health Study, and comparisons are made to other techniques for handling left-censoring as well as to estimates based on a traditional multi-state life table model.

An Evaluation of the Sensitivity of Multivariate Disease Monitoring Schemes to Parameter Estimation

✱Jeffrey M. Boone, The University of Alabama, Tuscaloosa, AL 35487, *jmboone1@bama.ua.edu*; Suhba Chakraborti, The University of Alabama; J. Michael Hardin, The University of Alabama

Key Words: statistical surveillance, multivariate surveillance, monitoring, parameter estimation, infectious disease

Multivariate surveillance schemes are useful in simultaneous monitoring of reporting units such as disease incidence from multiple geographic locations. Many of the currently available schemes assume that the models are specified correctly so that model parameters are assumed known. However, in real life applications, the model and the associated parameters must be estimated. This research focuses on evaluating the sensitivity of various multivariate monitoring schemes in order to detect incidents of interest when model parameters are estimated. Preliminary results, based on computations and simulations will be presented.

Individualized Absolute Risks from Epidemiologic Data

✱Anne S. Reiner, Memorial Sloan-Kettering Cancer Center, 307 East 63rd Street, Third Floor, New York, NY 10065, *reinera@mskcc.org*; Colin B. Begg, Memorial Sloan-Kettering Cancer Center; Elyn Riedel, Memorial Sloan-Kettering Cancer Center; Marinela Capanu, Memorial Sloan-Kettering Cancer Center; Jonine Bernstein, Memorial Sloan-Kettering Cancer Center

Key Words: absolute risk, contralateral, breast cancer, epidemiology

Relative risks in epidemiologic studies do not always have the clinical utility of individual absolute risks, which can be derived from epidemiologic data. The Women's Environmental Cancer and Radiation Epidemiology (WECARE) study has a unique design, where controls are women with unilateral breast cancer and cases are women with contralateral breast cancer (CBC). We calculate absolute risk of developing CBC by combining relative risks from a conditional logistic model from WECARE with baseline age-specific CBC rates from the general population using Surveillance Epidemiology and End Results (SEER) data. This method projects individualized risk estimates of developing CBC over a span of time and across a constellation of risk factors. This example highlights the method's clinical relevance, as women who survive a first primary breast cancer are at risk for CBC.

A Statistical Model for Real-Time Estimation of the Cumulative Confirmed Dengue

✱Pei-Hung Chuang, National Yang-Ming University, No.155, Sec. 2, Linong St., Beitou District, Taipei, International 112 Taiwan, *d49307013@ym.edu.tw*; I-Feng Lin, National Yang-Ming University; Jen-Hsiang Chuang, Centers for Disease Control and Prevention, Taiwan

Key Words: dengue fever, real-time estimation, syndromic surveillance system

An effective infectious disease surveillance system relies on a precise real-time cumulative number of cases. Since the laboratory confirmation of the case diagnosis is time-consuming, there is often a time-lag between the number of reported cases and the number of confirmed cases. In this study, we developed a statistical model to estimate the daily cumulative number of positive cases using dengue notification data. We assumed that the onset-to-confirmation time, the lag time between the date of disease onset and date of confirmed diagnosis, follows a Gamma distribution. Using the information such as reported date, onset date, confirmation date, and the diagnosis status of cases, we then calculated a conditional probability model to estimate the daily cumulative number of positive cases and the epidemic curve. Finally we evaluated the timeliness and the accuracy of this proposed method.

Lots of Data, Not Very Many People: Analysis of Multidimensional Profiles with Small Sample Sizes

✱Leann Myers, Tulane University, Dept. Biostatistics, SPHTM, 1440 Canal St., Suite 2001, New Orleans, LA 70001, *myersl@tulane.edu*; Margo A. Sidell, Tulane University; Michael J. Ferris, The Research Institute for Children

Key Words: multidimensional data, profile analysis, cluster analysis, simulations, multivariate analysis

Recent studies (e.g., Ferris, et al 2007) have quantitatively characterized the vaginal flora of normal women and those with bacterial vaginosis using PCR-amplification techniques. The number and diversity of the bacterial species inhabiting the vaginal epithelium create a data structure characterized by multiple, correlated outcomes. Because of the time-consuming nature of PCR methods, sample sizes are typically small. Given the large number of measures relative to the number of subjects, comparing or characterizing women on the basis of the bacterial profile presents an analytic challenge. We compared different approaches to analysis of these data, ranging from simple frequency distribution approaches to classical profile and cluster analyses to microarray analysis methods and simulations. We also demonstrate how the methods apply to data from other fields.

A Statistical Analysis of Historical Pandemic Influenza Data

✱Emily J. Powell, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, *emily.jpowell@yahoo.com*

This work started as a summer student project at Los Alamos National Laboratory involving using a simulation model to understand infrastructure impact of a disease outbreak. Inputs to the model included parameters associated with a pandemic influenza outbreak such as fatality rate and attack rate. The project goal was to explore ways to incorporate limited historical information to characterize possible distributions to assume for these inputs and have the student learn about application of Bayesian statistical methods and analysis tools. We model historical pandemic influenza data using a nested binomial model with beta priors that incorporate expert opinion and that is consistent with a simplified epidemiological model. We use a Bayesian inferential approach to obtain posterior distributions on case mortality rate and attack rate and study the effect of prior choices on these results.

Syndromic Surveillance Monitoring of Influenza Activity in Los Angeles County

✱Emily Kajita, Los Angeles County Department of Public Health, 313 N. Figueroa St. #222, Los Angeles, CA 90012, *ekajita@ph.lacounty.gov*; Akbar Sharip, Los Angeles County Department of Public Health; Patricia Araki, Los Angeles County Department of Public Health; Long Tai, Los Angeles County Department of Public Health; Bessie Hwang, Los Angeles County Department of Public Health

Key Words: Syndromic surveillance, Early event detection, Time series, Public Health, Disease trends

Los Angeles County (LAC) employs an early event detection system which analyzes emergency department (ED) visits for over 50% of the population. Visits are first classified into custom syndrome categories by searching chief complaint and diagnosis fields for key words, and then analyzed for statistical aberrations. Complementary systems include nurse hotline, coroner deaths, Reddinet© hospital volume, and over-the-counter medication sales data. Results are displayed on an internal website and monitored by analysts daily. System features that will be discussed, particularly with respect to monitoring influenza activity in LAC, include: 1) threshold calculations for detecting temporal aberrations and spatio-temporal SatScanT clusters, 2) forecasting counts using unobserved component models, and 3) contributions of complementary systems.

Estimating Risk for Transmission of Expanded Repeats Among Male Carriers of Intermediate Huntington Gene Alleles

✱Audrey E. Hendricks, Boston University School of Public Health, Boston University Medical Center, Department of Biostatistics, 715 Albany St., B617, Boston, MA 02118-2526, *baera@bu.edu*; Jeanne Latourelle, Boston University School of Medicine; Kathryn Lunetta, Boston University School of Public Health; Marcy MacDonald, Harvard Medical School; Adrienne Cupples, Boston University School of Public Health; James Gusella, Harvard Medical School; Richard Myers, Boston University School of Medicine

Key Words: Huntington, HD, genetic, probability, counseling, neurology

Huntington disease (HD) is a dominantly transmitted neurodegenerative disease that arises from an expansion of a CAG trinucleotide repeat on chromosome 4p16.3. CAG repeat allele lengths are defined as fully penetrant at = 40, reduced penetrance at 36-39, intermediate at 27-35, and normal at = 26. Fathers, but not mothers, with intermediate alleles (IA) are at risk of transmitting potentially penetrant HD alleles (= 36) to offspring. We estimated the conditional probability of a child having a penetrant allele given a father with an IA by applying Bayes and other probability rules to estimates of HD incidence, paternal birth rate, frequency of de novo HD, and frequency of IAs in the general population. The estimated risk that a male IA carrier will have a child with a penetrant allele ranges from 1/569 to 1/9480. These estimates may be useful in genetic counseling for male IA carriers.

Risk Factors Associated with Young Adults Nonmedical Prescription Drug Use (NMPDU) Using a National Sample: A Comparison of Recursive Partitioning Trees and Logistic Regressions

✱Lirong Zhao, University of Maryland School of Pharmacy, 220 Arch St, 12-606, Baltimore, MD 21201, *lzhao@rx.umaryland.edu*; Linda Simoni-Wastila, University of Maryland School of Pharmacy; Zhenqiu Liu, University of Maryland School of Medicine; Ming T. Tan, University of Maryland Greenebaum Cancer Center

Key Words: Risk factors, Outcome, Recursive partitioning, Logistic regressions

Objective: to compare the risk factors associated with NMPDU identified from logistic regression (LR) and recursive partitioning trees (RPT). Methods: 15,864 young adults (age 18-25) from the 2003 National Household Survey on Drug Abuse were used. 23 potential risk factors for social-demographic and behavior characteristics were extracted from data. LR (with stepwise variable selection) and RPT models were fit in SAS and R. Factors with P< 0.05 from LR and those that partitioned the sample in RPT were identified as risk factors. Results: 15.9% young adults used NMPDU in 2003. LR identified 12 risk factors including age, alcohol and cigar use, easy to obtain and use of several drugs, while RPT identified 6 of them. Conclusion: RPT is an alternative method to identify risk factors in health services and outcome research, especially when interactions among covariates are hard to determine.

Methods To Model the Impact of Respiratory Viruses on Asthma Risk

✱Pingsheng Wu, Vanderbilt Univrsity Medical Center, T1218 1161 21st Ave. South, Nashville, TN 37232, *wupingsheng@hotmail.com*; Tebeb Gebretsadik, Vanderbilt University Medical Center; William Dupont, Vanderbilt University Medical Center; Kecia Carroll, Vanderbilt University Medical Center; Marie Griffin, Vanderbilt University Medical Center; Tina Hartert, Vanderbilt University Medical Center

Key Words: childhood asthma, respiratory syncytial virus, rhinovirus, infancy viral infection

Different respiratory viral infections during infancy may carry a differential risk of childhood asthma. We determined the differential risk of asthma following infant respiratory syncytial virus (RSV) v. rhinovirus (RV) infection in infants born during 1995–2000 and continuously enrolled in the Tennessee Medicaid program who experienced at least one episode of clinically significant bronchiolitis. We used logistic regression model to differentiate the risk of asthma related to RV v. RSV bronchiolitis. 85% of infant bronchiolitis in August and September was attributable to RV, and 80% of bronchiolitis in December and January was attributable to RSV. Infants with August–September bronchiolitis were 42% more likely to develop asthma compared to infants with December–January bronchiolitis (95% CI: 1.24-1.64). Infancy RV-associated bronchiolitis is a strong predictor of childhood asthma.

Comparative Simulation Analysis of Modeling Heterogeneity in Closed Population Capture-Recapture Studies

✱Rob C. Wild, University of Oklahoma Health Sciences Center, Oklahoma City, OK , *robert-c-wild@ouhsc.edu*

Key Words: capture-recapture, epilepsy, log-linear models, latent class models, mixed logit models, bootstrap

Capture-recapture methods are used in epidemiology and public health to obtain estimates of population size. Individuals from the same population captured on more than one sample provide information for estimation of total population size. A real world application of the estimation of the number of individuals with epilepsy in a rural village in Burkina Faso, West Africa is used for motivation of a simulation of the effects of heterogeneous capture probabilities on population size estimates. Various scenarios using four lists and varying degrees of individual heterogeneity and sample overlap that “mimic” similar real world conditions are investigated through simulations in SAS/IML. The feasibility of marginal log-linear, latent class and mixed logit models under these conditions are addressed. Empirical bootstrapped confidence intervals are presented with population size estimates.

Oklahoma Tobacco Helpline: A Comparison of Sampling and Weighting Methods for Follow-Up Evaluation

✱Emily Leary, University of Oklahoma Health Sciences Center, *Emily-Leary@ouhsc.edu*; Bernardo Andrade, University of Oklahoma Health Sciences Center; Barbara Neas, University of Oklahoma Health Sciences Center; Laura Beebe, University of Oklahoma Health Sciences Center

Key Words: sampling, weighting, population simulation

Problem/Objective: To assess quit rates, follow-up surveys are given to a simple random sample of Oklahoma Tobacco Helpline (OTH) participants. We investigate different sampling and weighting methods to determine if there is a method that would yield better estimates for quit rates. Methods: Using registration data from fiscal year 2007, we simulate a population of OTH participants using R and sample this population using three different methods: simple random sampling, stratified, and probability proportional to size (PPS). Results: Preliminary data using SAS sampling methods indicate that PPS sampling produces the most accurate estimates, but the largest standard errors, for follow-up quit rates. Conclusions: Evaluating sampling methods for follow-up surveys can help determine better methods to estimate the number of tobacco users who have quit tobacco using OTH.

Geotemporal Spread of Influenza A in Canada and United States

✱Dena L. Schanzer, Public Health Agency of Canada, 100 Eglantine Driveway, Tunney's Pasture, AL:0602B, Ottawa, ON K1A 0K9 Canada, *dena_schanzer@phac-aspc.gc.ca*; Trevor Dummer, Dalhousie University; Joanne Langley, Dalhousie University; Samina Aziz, Public Health Agency of Canada; Brian Winchester, Public Health Agency of Canada; Theresa Tam, Public Health Agency of Canada

Key Words: public health policy, disease transmission, geo-temporal, mapping, influenza, epidemiology

A better understanding of the spread of annual influenza epidemics is considered key to pandemic planning. Based on weekly reports of laboratory confirmed influenza (FluWatch, Canada and CDC Flu, US) plus other Canadian data with finer spatial details, we characterized the geotemporal spread of influenza over the 1997/98-2005/06 period at various geographic scales. A composite epidemic curve was created by centering local epidemics relative to their epidemic mid-point. The epidemic growth rates as well as differences in regional timing were compared. Variation in direction of spread suggests that it is unlikely that travel alone is responsible for the dynamics of influenza transmission. Influenza appears to be geographically widespread early in the epidemic, and the annual epidemic likely starts independently in multiple regions. The epidemic starts in the fall though peaks in winter.

243 ASA College Stat Bowl II

ASA, ENAR, IMS, SSC, WNAR

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Round 2

✱Winners from Session 1, Six players will advance from Round 1 to Round 2

Round 2 will consist of two games. Six players will advance from Round 1 to Round 2, three players going to each game. The six players will be determined by the winners of the four games, plus the two non-winning players that score the highest. The winners of the two Round 2 games will face off in the Finals. Each player will receive team points depending upon how they finish (4 points for champion, 3 points for runner up, 2 points to semifinalists not advancing, 1 point to participants not advancing to round 2). Team accumulating the most points will be declared team champion. Ties for team totals are broken by the highest average score for all members of the teams in question.

244 Developments in Spatial/Environmental Health Modeling

Section on Statistics and the Environment, Section on Statistics in Epidemiology, WNAR

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Estimating a Pollutant's Policy-Related Background Level with Deterministic and Statistical Models

✱James V. Zidek, University of British Columbia, Department of Statistics, Vancouver, BC V6T 1Z2 Canada, *jim@stat.ubc.ca*; Zhong Liu, University of British Columbia; Nhu Le, BC Cancer Research Center

Key Words: air pollution, deterministic models, ozone, environmetrics, MAQSIP, policy related background

A PRB (Policy Related Background) level of a criterion pollutant such as ozone addressed in US regularity policy, is its (hypothetical) level if there were no anthropogenic sources in North America. It is a foundation above which to set regulatory standards but it cannot be measured directly so is inferred by deterministic CTMs (chemical transport models). The Bayesian melding approach of Fuentes and Raftery offers an approach for comparing CTM outputs with measurements to check accuracy. This paper describes the result of using that method along with aspects of the MCMC involved in this very parameter high dimensional parameter problem. Finally, some simpler alternative approaches to recalibrating the CTM outputs will be described.

Evaluating the Performance of Spatio-Temporal Bayesian Models in Disease Mapping

✱Maria Dolores Ugarte, Public University of Navarra, Statistics Department, Edificio Magnolios,Campus de Arrosadia, Pamplona, International 31006 Spain, *lola@unavarra.es*; Tomas Goicoa, Public University of Navarra ; Berta Ibañez, Fundación Vasca de Innovación e Investigación Sanitarias ; Ana F. Militino, Public University of Navarra

Key Words: Bayesian Decision Rules, Models Sensitivity and Specificity, bias and MSE of relative risks estimates, Hierarchical Bayesian Models

In this work, alternative Bayesian spatio-temporal models are fitted using MCMC techniques. The models performance is analyzed through a simulation study with two objectives in mind: the first one is to evaluate the relative bias and the relative MSE of the posterior mean relative risks, and the second one is to investigate recent Bayesian decision rules to detect raised-risk areas in a spatio-temporal context. The simulation study is based on colorectal cancer mortality data in males from Navarra, Spain at four five-year time windows. When there are a number of high-risk areas in some of the time periods we conclude that the bias of the posterior mean relative risks could be substantial. The decision rules to detect these high-risk areas should be determined with caution. A final rule combining alternative threshold and cut-off values for the different time periods seems to be needed.

Space-Time Environmental Interaction Modeling for Small-Area Health Outcomes

✱Andrew B. Lawson, University of South Carolina, Dept of Epoidemiology & Biostatistics, 800 Sumter Street, Columbia, SC 29208, *alawson@gwm.sc.edu*; Ahmed Al Hadrani, University of South Carolina

Key Words: spatial, temporal, mapping, disease, environmental, exposure

The analysis of space-time data is underdeveloped in general, and restricted to smoothing models for disease risk estimation in particular space-time applications. Convolution models with autoregressive temporal effects are common with interactions modeled as either independent or nonseparable. Often in environmental applications the focus is in estimation of regression parameters relating risk to exposures or their surrogates. Distance and direction functionals are important. However conventional convolution models fair badly when spatial aliasing of effects could be present. In addition, most previous models ignore the possibility of complex interactions between time and exposure effects. In this talk, we will examine a dynamic model for the linkage between putative source locations and outcomes. We also examine the possibility of modeling latent structures in the exposure time.

245 New Statistical and Computational Methods for Analysis of Genomic Data with a Graphical Structure ▲

Biometrics Section, Biopharmaceutical Section, WNAR
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Statistical Inference of Network Structure

✱Wing H. Wong, Stanford University, Department of Statistics, Stanford, CA 94305, *whwong@stanford.edu*

Key Words: network, Bayesian inference

Recently, genomics researchers have found Bayesian networks to be a useful approach to the modeling of data from the analysis of gene regulatory relations and signaling events. In this talk, we will review the statistical and computational challenges in this type of analysis. We will present preliminary results on how to handle missing data and how to search the space of network topology efficiently. We will also discuss the use of specialized hardware to handle some parts of the computation.

Pathways, Priors, Predictions: Network Inference in Cancer Systems Biology

✱Sachi N. Mukherjee, University of Warwick, Department of Statistics, Coventry, CV47AL United Kingdom, *S.N.Mukherjee@warwick.ac.uk*

Key Words: Systems biology, Graphical models, MCMC, Cancer biology

The Mitogen-Activated Protein Kinase or MAPK pathway is a biochemical pathway which plays a central role in cellular decision-making. Aberrant functioning of the pathway is heavily implicated in a number of cancers. Although the MAPK system is one of the best studied biological networks, cancer-specific pathway features remain poorly characterized, and relatively little is known regarding isoform-specific connectivity. I will present a statistical approach to characterizing pathway connectivity in this setting, using a combination of biochemical data and biological prior knowledge. I will discuss some of the methodological aspects of the problem and go on to present results obtained from breast cancer data. I will also discuss some of the caveats of this type of analysis and highlight the rich opportunities and challenges for statistical research in cancer systems biology.

Use of the Graph Laplacian To Analyze Network Data

✱Elizabeth Purdom, University of California, Berkeley, 367 Evans Hall #3860, Berkeley, CA 94720-3860, *epurdom@stat.berkeley.edu*

Key Words: networks, graphs, laplacian, high dimensional data

Many biological processes are represented by graphs or networks. Incorporating information from these graphs into an analysis of the experimental data is generating increasing interest. Results from graph theory suggest using the spectral decomposition of the adjacency matrix (or certain transformations of it) in order to find interesting features of the numerical data that also reflect the structure of the graph. The adjacency matrix representation of graphs also underlies similar kernel methods that have been used to jointly analyze graphs and data. Because of the relative numerical simplicity of manipulating the adjacency matrix, this approach is particularly useful in a large scale analysis, such as relating microarray experiments with a database of biological networks. We also describe adaptations to this approach to make the analysis more sensitive to biological networks.

Bayesian Variable Selection in Structured High-Dimensional Covariate Spaces, with Applications in Genomics

✱Nancy R. Zhang, Stanford University, Sequoia Hall, 390 Serra Mall, Stanford, CA 94305-4065, *nzhang@stanford.edu*; Fan Li, Harvard Medical School

Key Words: Bayesian variable selection, Ising model, DNA copy number data, motif analysis, Markov chain Monte carlo

We consider the problem of regression modeling in high dimensional spaces where there is known structure among the covariates. Such problems are becoming increasingly relevant as high-throughput data collection schemes become increasingly common. Our approach builds on the framework of Bayesian variable selection, where we formulate a general Ising prior on the model space for incorporating structural information. However, certain computational and statistical problems arise that are unique to such high dimensional, structured settings, the most interesting being the phenomenon of phase transitions. We propose theoretical and computational schemes to mitigate these problems. As examples we discuss two specific applications in genomics, one arising from DNA copy number analysis, and the other arising from the modeling of transcription factor binding sites in DNA sequences.

246 Confidentiality Training: What Does It Mean, and How Is It Practiced? ▲

Committee on Privacy and Confidentiality, Section on Statistics in Epidemiology, Social Statistics Section
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Required Confidentiality Training: What Does It Mean and How Is It Practiced? An Overview of Concepts, Practices, and Key Considerations

✱Alvan O. Zarate, National Center for Health Statistics, 33 Toledo Road Rm 7116, Hyattsville, MD 20782, *aoz1@cdc.gov*

Key Words: Confidentiality, Training, Privacy, Disclosure

This paper will set the stage for an in-depth discussion of existing concepts and approaches to confidentiality training in major sectors of the statistical community. Major variants of key concepts (confidentiality, privacy, statistical disclosure) will be discussed followed by an overview of features of existing training schemes. Additional considerations will be offered regarding: training targets (staff, data users, students, statisticians/non-statisticians, data base administrators, general public); the training providers (ASA/other professional societies, federal government, private sector firms); the components of training (on-line courses, seminars,); training timing and frequency (orientation, continuing) and minimum content.

Confidentiality Training in a Federal Statistical Agency

✱Christa D. Jones, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *christa.d.jones@census.gov*

Key Words: confidentiality, training, federal statistical agency, data stewardship

The U.S. Census Bureau is the largest federal statistical agency and its mission is to serve as the leading source of information about the nation's economy and its people depends on our ability to gain public trust and cooperation. Confidentiality is a promise and commitment we make to respondents. Training is the first step in ensuring this commitment. The Census Bureau has recently updated its core confidentiality training effort, the annual Title 13 training, and is looking to expand its training efforts. Whether training takes the form of computer based learning, town halls, or specific programmatic efforts, it is increasingly viewed as important opportunity to strengthen the data stewardship culture within the agency.

Implementing CIPSEA in a Brand New Statistical Unit: Implications for Staff, Contractor, and Agent Training

✱Jonaki Bose, Substance Abuse and Mental Health Services Administration, Division of Population Surveys, Office of Applied Studies, 1 Choke Cherry Road, Room 7-1014, Rockville, MD 20857, *Jonaki.Bose@samhsa.hhs.gov*

Key Words: Statistical Unit, CIPSEA Training, OAS

The Office of Applied Studies (OAS) in the Department of Health and Human Services, Substance Abuse and Mental Health Administration was designated as a statistical unit under CIPSEA by the Office of Management and Budget (OMB) in November 2006. In June 2007, OMB released guidelines on the implementation of CIPSEA. CIPSEA provides strong protections for data and the ability to designate agents in order to share restricted-use data. When sharing restricted-use data with agents, there are a number of responsibilities that include agent confidentiality training and agent monitoring. This paper addresses how OAS has developed and implemented the annual training and certification requirements in CIPSEA among its staff, contractors and other agents.

Confidentiality Training for the National Opinion Research Center's Data Enclave

✱Julia Lane, National Science Foundation, 4201 Wilson Boulevard, 905 N, Arlington, VA 22230, *jlane@nsf.gov*; Timothy M. Mulcahy, National Opinion Research Center; Pascal Heus, Open Data Foundation

Key Words: Confidentiality, Training, Collaboratory, Virtual user community, Data Enclave, Microdata access

The NORC training course has four aims. The first is to provide authorized researchers with an understanding of the legal and ethical context within which they are being provided access to the data, together with the practical implications—notably an overview of the basics of disclosure protection. The second is to acquaint researchers with the logistics of working in the enclave from their offices, developing metadata "best practices" and how to work in a geographically distributed collaborative environment. The third is to provide training in the various aspects of the data with which they'll be working, and is typically provided by the data producer. The final component is to develop personal contacts among the researchers, with personal relationship, which research has shown is critical to developing a successful virtual user community.

247 Nonparametric Bayesian Data Analysis ●

Council of Chapters, Section on Bayesian Statistical Science, Section on Nonparametric Statistics

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Center-Adjusted Inference for a Nonparametric Bayesian Random Effect Distribution

✱Yisheng Li, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Unit 447, Houston, TX 77030, *ysli@mdanderson.org*; Peter Mueller, M.D. Anderson Cancer Center; Xihong Lin, Harvard University

Key Words: Fixed effects, Generalized linear mixed model, Moments of a Dirichlet process, Post processing, Random effects, Random probability measure

Dirichlet process (DP) prior models are a popular choice for semiparametric Bayesian general or generalized linear mixed models. The base measure of the DP prior for the random effect distribution is often specified as a normal distribution with a zero mean. The fact that this prior allows the random effect distribution to have a nonzero mean complicates the interpretation of inference for the fixed effects paired with the random effects. The resulting inference for the fixed effects may be poor. We propose a post-processing technique to adjust the inference. The approach uses a parametrization of the DP with a base measure centered at an unknown nonzero mean. The proposed adjustment for the analytically evaluated moments of the DP can be conveniently incorporated into the MCMC simulations. We evaluate the performance of the method through simulations and apply it to a real data set.

Sparse Modeling of Conditional Response Distributions with Many Predictors

✱David Dunson, National Institute of Environmental Health Science, 111 T. W. Alexander Dr., Research Triangle Park, NC 27709, *dunson1@niehs.nih.gov*

Key Words: Nonparametric Bayes, Variable selection, Density regression, Shrinkage, Quantile regression, Large p, small n

There is an increasingly rich literature on methods for accommodating high-dimensional predictors in regression, either to identify promising subsets of important predictors or for prediction. Such methods almost always focus on a mean regression setting in which the predictors are related to the mean of the response variable, with the residual distribution assumed to be constant or possibly changing according to a parametric model. Our focus is on using nonparametric Bayes methods to flexibly model the conditional response distribution, allowing the different quantiles of the distribution to change differentially with a high-dimensional set of candidate predictors. We propose an approach based on the kernel stick-breaking process, and illustrate the method with epidemiologic applications.

Bayesian Semiparametric Modeling and Inference for Longitudinal Diagnostic Testing Data

✱Wesley O. Johnson, University of California, Irvine, 2322 Bren Hall, Irvine, CA 92697-1250, *wjohnson@ics.uci.edu*; Michelle Norris, University of California, Davis

Key Words: Bayesian Nonparametric, Diagnostic testing, Longitudinal, Dirichlet Process Mixture, Change point

We propose a model for joint longitudinal data where one response variable is continuous and the other is categorical; both are diagnostic measures for infection. Individuals may or may not become infected during the course of the study. We include a latent time varying infection status variable for each individual. Both response variables are modeled to behave differently after infection than before. Classification is a primary goal. It is anticipated that there will be clusters of individuals with distinct response patterns in the continuous variate through time and so a Dirichlet Process Mixture model is employed to model random effects and/or functional random effects for each individual. Inferences are based on MCMC methods for varying-dimension parameter spaces. Methods are applied to ELISA and fecal culture screening data for Johne's Disease in cattle.

248 Controversy and Convergence on Modeling Synergism in Drug Combination Studies ●▲

ENAR, Biopharmaceutical Section, WNAR, Biometrics Section

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Trial by Fire: Comparison of Statistical Models for Studying Drug Combinations

✱William R. Greco, University at Buffalo, SUNY, School of Pharmacy & Pharmaceutical Sciences, Buffalo, NY 14260, *rosgreco@buffalo.edu*

Key Words: synergy, response surface, antagonism, additivity

To reveal common threads and dissimilar elements among rival modeling approaches for binary drug combinations, 16 pharmacometric modelers were each invited to analyze two previously-published [Faessel, et al,. Cancer Res. 58:3036-3050, 1998] data sets which came from the study of two anticancer agents against cultured cells in a 96-well plate cell growth assay. The first data set includes the complexity of local regions of synergy and antagonism on the same response surface; and the second data set includes the rare finding of extreme (super) synergy. Each of these two large data sets came from two 5-plate stacks (880 randomly treated wells per experiment), with a fixed-ratio design. The data sets are rich both in design points and in replicates, facilitating innovative modeling approaches. The results of the rival approaches will be summarized, compared and evaluated.

Application of Mixed Models in Drug Combinations

✱Maiying Kong, University of Louisville, Department of Bioinformatics and Biostatistics, University of Louisville, Louisville, KY 40292, *maiying.kong@louisville.edu*; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center

Key Words: Assess drug interaction, Loewe additivity model, Drug combinations

Studies of interactions among biologically active agents have become increasingly important in many branches of biomedical research. However, in many in vivo experiments, the effects measured are not only related to the combination doses but also the experimental conditions. For example, when several trays with 96 wells are used for the same experiment, the numbers of cells may be added differently due to different persons, and the sitting time or temperature may be different during different days, thus resulting different reading values even though the combination doses are the same. In this talk, we use random effect models with drug effect being treated as fixed effects, and experimental condition being treated as random effects. Thus, we can make more precise inferences for assessing drug interactions.

Controversy and Convergence on Modeling Synergism in Drug Combination Studies

✱Ming T. Tan, University of Maryland Greenebaum Cancer Center, MSTF 2nd Floor, Room 261, 10 S. Pine Street, Baltimore, MD 21201, *mttan@som.umaryland.edu*; Hong-Bin Fang, University of Maryland Greenebaum Cancer Center

Key Words: Drug Combination, Synergy, Experimental design, Interaction Index, Loewe Additivity, Sample size

Preclinical experiment on multi-drug combination has an increasingly important role in (especially cancer) drug development. We propose a general method based on the uniform scattered points in the experimental domain for dose finding and sample size determination for combination studies of multiple drugs in a semi-parametric statistical model applicable to both in vivo and in vitro experiments. The design maximizes the power to detect departure from additivity. We demonstrate the application of the method to a study on SAHA and Ara-C where subsequent data collected according to this design have identified synergistic combinations that would been missed with more classic designs. We also validate the design utilizing the combinations of two enzyme inhibitors where comprehensive data are available on the dose response.

249 Would the Real Data Please Stand Up? Data Comparisons for Policy Analysis ●▲

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section, Scientific and Public Affairs Advisory Committee

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Measuring Income and Poverty in Four Surveys: An Overview

✱Connie Citro, Committee on National Statistics, 500 Fifth Street, NW, Washington, DC 20001, *ccitro@nas.edu*; Joan L. Turek, U.S. Department of Health and Human Services; Gabrielle Denmead, Denmead Services and Consulting; Sameer Desale, Synectics for Management Decisions Inc.; Brian James, U.S. Department of Health and Human Services

Key Words: income, poverty, family structure, policy analysis, program benefits, health insurance coverage

Policymakers use national surveys to paint a picture of the U.S. population along a variety of dimensions such as poverty status, receipt of program benefits, demographic characteristics and health insurance coverage. Inferences are drawn about need and eligibility for federal programs based on estimates produced by these surveys. Joan Turek will present the findings of research by Gabrielle Denmead, Joan Turek, Brian James, and Sameer

Desale that develops comparable measures of income, family structure, and poverty across four surveys. It examines whether the same picture of the U.S. population is presented by these surveys. The four surveys are the Annual Social and Demographic Supplement to the Current Population Survey, the Medical Expenditure Panel Survey Household Component, the National Health Interview Survey, and the Survey of Income and Program Participation.

How Much Is It Worth: Comparing the Policy-Analytic Value of Wealth Data from Four Surveys

✱John L. Czajka, Mathematica Policy Research, Inc., 600 Maryland Avenue S.W., Suite 550, Washington, DC 20024, *jczajka@mathematica-mpr.com*

Key Words: assets, wealth, debts, policy analysis, program eligibility

Data on the value of people's assets and debts are particularly difficult to collect because respondents often cannot provide balances without consulting records, which may not be readily accessible. Moreover, the value of certain major assets, such as a home, property or business, is not typically available from any records. These factors are compounded by the sensitive nature of wealth holdings, resulting in high item nonresponse and substantial error in reported values. Some policy applications of wealth data—assessing program eligibility—require greater accuracy than the data are able to provide. This paper examines the measurement of wealth in four surveys: Survey of Income and Program Participation, Survey of Consumer Finances, Panel Study of Income Dynamics and the Health and Retirement Survey. Estimates are presented for the entire and policy relevant subpopulations.

Misreporting of SSI and Social Security Benefits: Evidence from Comparisons of Survey and Administrative Data

Paul Davies, Social Security Administration; ✱Lynn Fisher, Social Security Adminsitration, 9th Floor ITC Building, 500 E St. SW, Washington, DC 20254, *Lynn.Fisher@ssa.gov*

Key Words: Supplemental Security Income recipients, Social Security benficiaries, asset income, poverty, well-being, administrative data match

This paper presents findings on the misreporting of Supplemental Security Income (SSI) and Social Security benefits in the Survey of Income and Program Participation (SIPP) and the March Current Population Survey (CPS). Using Social Security administrative data matched to SIPP and CPS, several analyses conducted by the Social Security Administration's Office of Policy have identified substantial reporting error in SSI participation and benefit amounts. One common error is for respondents to misreport SSI as Social Security Income. Such misreporting has implications for the measurement of poverty and well-being among SSI recipients and Social Security beneficiaries. Combined with under-reporting of asset income and difficulties in the measurement of defined contribution pension income, this misreporting also has implications for estimates of elderly benefits.

A Comparison of the Health Insurance Coverage Estimates from Four National Surveys and Six State Surveys

✱Michael Davern, University of Minnesota/SHADAC, 2221 University Avenue, suite 345, Minneapolis, MN 55110, *daver004@umn.edu*

Key Words: health insurance coverage, uninsured

Health insurance coverage in the U.S. and the characteristics of the uninsured population are of substantial importance because health coverage is related to individual health. Researchers have developed many explanations for why the national and state health insurance estimates from the Current Population Survey's Annual Social and Economic Supplement (CPS-ASCE) differ from state-initiated survey data collection activities, the National Health Interview Survey, the Medical Expenditure Panel Survey Household Component and the Survey of Income and Program Participation. We present results from a systematic comparison of health insurance estimates from these surveys to the CPS-ASCE using bivariate and multivariate techniques to see where the largest pockets of difference among these surveys is coming from.

250 Statistical Challenges in Online Advertising and Search ●▲

Section on Statistical Computing

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Understanding Online Advertisers

✱Daryl Pregibon, Google, Inc., 76 Ninth Ave, New York, NY 10011, *daryl@google.com*; Diane Lambert, Google, Inc.

Key Words: selection bias, internet search, network models, ad effectiveness

Users of search engines (SEs) are often regarded as the primary constituents of successful SEs. The paying customers, namely advertisers, enable the free services that SEs provide and yet they are not as well understood as the user population. We discuss several studies that focus on advertisers: methods for evaluating usefulness of new tools to manage their campaigns, and methods for assessing campaign effectiveness. We also introduce the concept of advertiser graphs as a vehicle to better understand advertisers and their inter-relationships.

Placing Online Advertisements Based on Context

✱Kishore A. Papineni, Yahoo! Research, 110 W 40th St, New York, NY 10018, *kpapi@yahoo-inc.com*

Key Words: online advertisements, pay-per-click, importance sampling, click prediction

We examine placing pay-per-click advertisements on web pages. This problem involves matching the context (user and page content) with a small set of advertisements. The opportunity to place the advertisements may be auctioned online. To participate in such an auction, the placement mechanism must evaluate the expected revenue for a slate of advertisements for each user-page context. This involves predicting rare events (clicks). Maximizing revenue from clicks involves depleting the (finite) budgets of the advertisements based on forecasts of matching contexts for each advertisement. For this forecasting to be feasible, contexts may be classified into a hierarchy. Training a predictor of clicks based on past clicks involves sampling negative examples and examining billions of predicates. This talk examines some statistical aspects that underlie this problem.

Social and Semantic Structures in Web Search

✱Andrew Tomkins, Yahoo! Inc., 701 First Ave, Sunnyvale, CA 94089, *atomkins@yahoo-inc.com*

Key Words: Web search

Nonprofessional creation of public online content has outstripped professional content creation of all forms, both online and offline. However, two orders of magnitude more content are created daily to flow through social networks, with as much as two more orders of magnitude still to come as user engagement increases. At the same time, content is diversifying in creation, consumption, and nature. In this talk, I'll cover these trends and their implications. I'll present some research results regarding online communities and semantic structures, and will close with some challenges for the future.

251 Signal Classification and Pattern Discovery ●

ASA Interest Group on Statistical Learning and Data Mining, Section on Nonparametric Statistics, Section on Statisticians in Defense and National Security
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

A VAR Model for Effective Connectivity in Slow Event-Related fMRI Designs

✱Wesley Thompson, University of Pittsburgh, 6476 Monitor Street, Pittsburgh, PA 15217, *wesleyt@pitt.edu*

Key Words: Functional Data, fMRI, Effective Connectivity, Vector Autoregression, Smoothing, Dynamic State Space Model

Researchers have become increasingly interested in exploring dynamic, directed relationships (effective connectivity) among brain regions involved in processing a given stimulus or mental task. An experimental paradigm commonly used for determining effective connectivity in fMRI studies is the slow event-related design, which presents stimuli at sufficient temporal spacing for determining within-trial response trajectories of BOLD activation across regions. We will discuss an adaptation of vector autoregressive models tailored to effective connectivity analysis of nonstationary data from slow event-related fMRI designs. Connectivity coefficients have been estimated via a tensor product of basis expansions and allowed to vary randomly across subjects.

Detection of Cognitive Fatigue from EEG Signals

✱Raquel Prado, University of California, Santa Cruz, Baskin Engineering UCSC, AMS, 1156 High St. MS:SOE2, Santa Cruz, CA 95064, *raquel@ams.ucsc.edu*

Key Words: cognitive fatigue, time series, multi-processes, structured priors, sequential updating

Automatic detection of early signs of mental fatigue is key for increasing safety and human performance in vehicle operation, air traffic control and space missions. In this work we study electroencephalographic signals (EEGs) recorded in subjects who performed continuous mental arithmetic for up to 180 minutes. The main goal of the analysis is online detection of cognitive fatigue from the EEGs. We model various levels of alertness via mixtures of autoregressive processes. In particular, we propose the use of structured prior distributions on functions of the model parameters that are scientifically interpretable. We develop and illustrate the performance of algorithms for sequential posterior estimation of the model parameters over time. This approach allows us to provide practitioners with a probabilistic assessment of cognitive fatigue over time.

Discovering Patterns of Connectivity Within the Human Brain

✱F. DuBois Bowman, Emory University, Department of Biostatistics, 1518 Clifton Road, N.E., Atlanta, GA 30322, *dbowma3@sph.emory.edu*

Key Words: fMRI, neuroimaging, connectivity, classification, DTI

There has been recent interest in the neuroimaging community in investigating connectivity of the human brain. One common approach is to evaluate functional connectivity, which involves a search for correlations in brain activity measurements between spatially distinct regions. Given the intricate system of intercommunicating neuronal networks via white matter fiber tracts, it is likely that functional connections between two brain regions are mediated by anatomical connections. Diffusion tensor imaging (DTI), which enables the measurement of restricted water diffusion in tissue, provides a means for constructing probable white matter anatomical tracts. In this talk, we discuss statistical methods for determining functional connectivity, and we explore methods for combining functional and anatomical information derived from DTI.

252 Consulting Service Models for Academic Medical Centers: What Works, What Doesn't Work, Reaching Out, and Filling the Growing Need ●▲

Section on Statistical Consulting, WNAR, Section on Statistics in Epidemiology
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

The Staff Model, the Independent Contractor Model, and Marketing the Consulting Service in a Teaching and Research Hospital

✱Shelley Hurwitz, Brigham & Women's Hospital, Harvard Medical School, 1620 Tremont St., CCI/BWH-3, Boston, MA 02120, *hurwitz@hms.harvard.edu*

Key Words: consulting, academic, medical

There is increasing acknowledgment of the value of statistics in medical research by investigators and the public. Consequently the demand for statistical services is growing. The presenter designed and directed two consulting services at an academic center associated with Harvard Medical School, with different operating models and fee structures. One had a large number of independent consultants over six years. The other had a smaller number of statisticians on staff for four years so far. The features that work well and the features that do not work so well will be described. In line with the theme for JSM 2008, the focus will be on educating investigators and research teams, and reaching out to those in need of statistical support through marketing. As a result, we fill the increasing need, promote the necessity of statistical rigor, and improve our profession's image.

Statistical Consulting for a Newly Formed Department of Biostatistics in a Highly Collaborative Environment at the University of Kansas Medical Center

✱Jonathan Mahnken, The University of Kansas Medical Center, 3901 Rainbow Blvd., MSN 1026, Kansas City, KS 66160, *jmahnken@kumc.edu*

Key Words: Consulting metrics

The foremost role of statisticians in our academic medical center has been to provide collaborative support to the research enterprise. Over the past few years, many of the statisticians on campus were affiliated with a consulting unit within the Dept. of Preventive Medicine. Very recently, this unit has grown to a campus-wide center and, in 2007, to a Dept. of Biostatistics within our School of Medicine. Tracking support provided to grants and projects played a key role in providing metrics of productivity to our campus administrators to justify a Dept. of Biostatistics. This talk will present the types of support provided throughout our medical center, the means of obtaining support, payment options for investigators, metrics used to evaluate our collaborative productivity, and time allocation between consulting and other responsibilities of faculty biostatisticians in our department.

Consulting and Collaboration Model for a Health Sciences Center

✱Guy Brock, University of Louisville, 555 S Floyd St, Louisville, KY 40202, *guy.brock@louisville.edu*; Rudolph Parrish, University of Louisville

Key Words: consulting, collaboration, center grants, public health, statistical education

In this talk, I will present our current framework and vision for the Statistical Consulting Center (StCC) at the University of Louisville. Our StCC primarily serves the needs of four schools within the Health Sciences Center (Public Health, Nursing, Medicine, and Dentistry). The current model is a mix of consulting services provided on a fee-for-service basis, and collaborative enterprises. I will present our vision of the StCC as a collaborative center, with specialty subdivisions to address the diverse statistical needs of the HSC. Topics addressed include incorporation of students on consulting projects, addressing the needs of NIH center grants, and collaboration versus consultation.

Consult, Coordinate, and Critique: How the University of Minnesota Department of Family Medicine Came To Be Ranked #2

✱Bruce A. Center, The University of Minnesota, Dept Family Medicine, 420 Delaware St SE, Minneapolis, MN 55455, *cente001@umn.edu*; Anne M. Weber-Main, The University of Minnesota; Carole J. Bland, The University of Minnesota

Key Words: Effective Grant proposals, Family Medicine, Statistical support in Academic medical departments, Collaboration

Statisticians are necessary and powerful partners during a research proposal's development, but our effectiveness is mediated by knowledge of the investigator's content area and time to guide them through multiple cycles of writing-review-rewriting. We describe how the U of MN Family Medicine Dept. used a core support team of three PhDs—with expertise in statistics, design, evaluation, measurement, and scientific writing/editing—to rise to #2 in the department's NIH ranking, without hiring more research faculty. In brief: 1) Investigators consult with us early to refine the project's aims, hypotheses, importance, design, and analytic plan. 2) We coordinate a timeline for the proposal's development by investigators and support staff. 3) We critique proposal drafts through formal internal peer review. We depict the underlying system: what worked, what didn't, and current status.

Teaching Statistics in the Consulting Environment: Experiences at the University of Iowa Biostatistics Consulting Center

✱M. Bridget Zimmerman, The University of Iowa, Dept of Biostatistics C22-E GH, 200 Hawkins Drive, Iowa City, IA 52242-1009, *bridget-zimmerman@uiowa.edu*

Key Words: Consulting, Teaching

The consulting environment provides opportunities to educate our clients on the important contributions of statistics in the research endeavor. About one-third of the clients served by the Biostatistics Consulting Center (BCC) are medical or graduate students, medical residents, and post-doctoral fellows working on an independent research project. For most of them, their statistical knowledge comes from one introductory statistics course that they had taken as an undergraduate or in medical school. The challenge with these interactions is how to engage the clients to think of statistics beyond the "p-value." In addition, we will also describe the role of the BCC in training biostatistics students. This talk will present examples in various areas at which teaching during the consulting process takes place.

253 Inference of Structure and Information Flow in Networks

IMS, Section on Statisticians in Defense and National Security

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Bayesian Inference for Stochastic Biochemical Network Parameters via a Diffusion Approximation

✱Andrew Golightly, Newcastle University, School of Maths & Stats, Herschel Building, Newcastle Upon Tyne, NE1 7RU United Kingdom, *a.golightly@ncl.ac.uk*

Key Words: Nonlinear diffusion process, Bayesian inference, Markov chain Monte Carlo

This talk considers stochastic differential equations for Systems Biology models derived from the Chemical Langevin Equation (CLE). After outlining the derivation of such models, Bayesian inference for the parameters is considered, based on state-of-the-art Markov chain Monte Carlo algorithms. Starting with a basic scheme for models observed perfectly, but discretely in time, problems with standard schemes and their solutions are discussed. Extensions of these schemes to partial observation and observations subject to measurement error are also considered. Finally, the techniques are demonstrated in the context of a simple stochastic kinetic model of a genetic regulatory network.

Optimal Designs for Network Traffic Measurement Schemes

George Michailidis, The University of Michigan; ✱Harsh Singhal, Unverisity of Michigan, 1929 Plymouth Road, #4038, Ann Arbor, MI 48105, *singhal@umich.edu*

Computer network traffic can be sampled at routers to determine the origin and destination of different packets. This is in turn used to estimate the volume of traffic between all origin-destination pairs. Use of higher sampling rate leads to better estimates but uses more resources. In this work we formulate the problem of determining the network-wide sampling rates as an optimal design problem. This problem is solved using semi-definite programming and simulations show that this design performs better than conventional sampling schemes.

Network Inference from Co-Occurrences

✱Robert Nowak, University of Wisconsin-Madison, 1415 Engineering Drive, Madison, WI 53706, *nowak@engr.wisc.edu*

Key Words: networks, monte carlo, genomics, bioinformatics, markov models, network dynamics

The discovery of networks is a fundamental problem arising in numerous fields of science and technology, including communication systems, biology, sociology and neuroscience. Unfortunately, it is often difficult to obtain data that directly reveal network structure, and so one must infer a network from incomplete data. This paper considers inferring network structure and transmission pathways from "co-occurrence" data; observations that identify which network components carry each transmission but do not indicate the order in which they handle the transmission. A maximum likelihood (ML) estimator based on a Markovian model of the network dynamics is proposed. For large-scale network inference, we propose a polynomial-time Monte Carlo algorithm, which is guaranteed to converge to a local ML point with high probability.

254 The SAMSI Program on Risk Analysis, Extreme Events, and Decision Theory

Section on Risk Analysis

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Extreme Co-Movements and Extreme Impacts in High-Frequency Data in Finance

✱Zhengjun Zhang, University of Wisconsin, 1300 University Ave., MSC 1220 Dept. Stat., Madison, WI 53706, *zjz@stat.wisc.edu*

Key Words: High frequency data, Exchange rate, Extreme dependence

Extreme co-movement and extreme impact problems are inherently stochastic control problems, since they will influence the decision taken today and ultimately influence a decision taken in the future. Extreme co-movements among financial assets have been reported in the literature. However, extreme impacts have not been carefully studied yet. In this paper, we use the newly developed methodology to further explore extreme co-movements and extreme impacts in financial market. Particularly, two FX spot rates are studied. Based on the results of our analysis with FX returns, we conclude that there exist extreme co-movements and extreme impacts in FX returns and care has to be taken when we employ portfolio optimization models, especially models without the ability of handling extreme dependencies.

Modelling Multivariate Extreme Dependence

✱Xiao Qin, Beihang University; University of North Carolina, Beijing; Chapel Hill, *miniqin@126.com*; Richard L. Smith, The University of North Carolina at Chapel Hill; Ruoen Ren, Beihang University

Key Words: Copula, Coefficient of tail dependence, Extreme value theory, Asymptotic independence, Asymptotic dependence

Ledford and Tawn (1997, JRSSB) propose a bivariate framework to model the joint tail distribution which incorporates positively and negatively associated asymptotical independence into asymptotical dependence and exact independence, and thus greatly extends the traditional approach based on bivariate extreme value theory. Their approach, however, is not general enough for dimensions higher than two, whereas the considerable complexity of real data implies the urgent necessity to model the tail dependence in higher dimensions. This paper generalizes their methodology into a multivariate context and suggests the relevant inference strategies. Both simulation and practical examples demonstrate the practical value of our extension.

Multivariate Analyses of Extremes

Luis R. Pericchi, University of Puerto Rico, Rio Piedras; Beatriz Mendes, Universidade Federal de Rio de Janeiro; Scott Sisson, University New South Wales; ✱Abel Rodriguez, University of California, Santa Clara, Santa Clara, CA, *abel@cesma.usb.ve*

Key Words: Extremes, Multivariate, Hefernan and Tawn, Hydrology

We model multivariate hydrological extremes (5-dimensional or more) focusing on data from rivers in the Caribbean and Australia. The model by Hefernan and Tawn (2004, JRSSB) and elaborations, are the main tools for the applications. Complex conditional multivariate inferences are shown to be possible under this model and its extensions.

Downscaling Extremes: A Comparison of Extreme Value Distributions in Point-Source and Gridded Precipitation Data

✱Elizabeth C. Shamseldin, University of North Carolina, Department of Statistics and Operations Research, Chapel Hill, NC 27599-3260, *shamseld@email.unc.edu*; Richard L. Smith, The University of North Carolina at Chapel Hill; Stephan Sain, National Center for Atmospheric Research; Dan Cooley, Colorado State University; Linda O. Mearns, National Center for Atmospheric Research

Key Words: Projections of Extreme Events, Re-analysis, Generalized Extreme Value Distribution, Generalized Pareto Distribution

Empirical and climatological evidence show precipitation extremes have increased during the 20th century. Understanding this is limited by spatial scaling: most evidence of past trends comes from rain-gauge data whereas future trends produced by climate models rely on gridded aggregates. We fit the GEV distribution to extreme rain-gauge and gridded events. The rain-gauge data are from monitoring stations and the gridded data from a re-analysis model (NCEP) and NCAR's Community Climate System Model. The results confirm that return values from rain-gauge data are typically higher than those from gridded data. The main contribution is the development of a family of regression relationships that also take spatial variations into account. Based on these results, we believe it is possible to project future changes in precipitation extremes at the station level based on climate model data.

Hurricanes and Global Warming

✱Richard L. Smith, The University of North Carolina at Chapel Hill, Dept of STOR, Chapel Hill, NC 27599-3260, *rls@email.unc.edu*; Evangelos Evangelou, University of North Carolina; Gabriel A. Vecchi, Geophysical Fluid Dynamics Laboratory; Thomas R. Knutson, Geophysical Fluid Dynamics Laboratory

Key Words: Bivariate time series, Environmental trends

There is an extensive literature documenting the increased incidence of Atlantic tropical cyclones and hurricanes over the latter half of the twentieth century, and their correlation with sea surface temperatures (SSTs). Tying these two facts together, it is widely believed that there is a direct correlation between increased hurricanes and human-induced global warming. However this association is by no means so clear-cut if the period of observation is extended back into the late nineteenth century (Vecchia and Knutson, 2007). In this paper, we consider bivariate time series models for the joint distribution of hurricane or tropical cyclone counts and SSTs. The results show that after taking account of common cyclical effects in the two series, there remains an increasing trend over the time period of study, but the evidence of a causal anthropogenic effect is still uncertain.

255 Survival Analysis Methods ●▲

Biometrics Section, Biopharmaceutical Section, WNAR

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Intent-to-Treat Duration of Response: Statistical Issues in Study Design and Analysis

✱Ritwik Sinha, Bristol-Myers Squibb Company, 5 Research Parkway, Research & Development, Wallingford, CT 06492, *ritwik.sinha@bms.com*; Tai-Tsang Chen, Bristol-Myers Squibb Company; Chao Zhu, Bristol-Myers Squibb Company

Key Words: duration of response, survival, ITT, sample size

The durability of response is increasingly used as an endpoint for cancer clinical trials, especially in some chronic indications, such as leukemia, in which the patients are expected to be in remission for a long period of time before progression or death. It is of clinical interest to compare the duration of response between two treatment arms among these patients. The traditional analysis of duration of response by definition can only include responders, since the time of origin, i.e., date of response, is only available among these patients. We propose an approach to determine the sample size when the primary endpoint of interest is duration of response. The proposed method is assessed using Monte Carlo simulation.

Use of Classification Tree Techniques To Determine the Likelihood of Pretransplant Death Among Patients Awaiting Liver Transplantation

✱Chung-Chou Chang, University of Pittsburgh, 200 Meyran Ave., Suite 200, Pittsburgh, PA 15213, *changj@pitt.edu*; Fiona M. Callaghan, University of Pittsburgh

Key Words: Classification tree, Competing risks

Scores derived from the model for end-stage liver disease (MELD) are currently used to determine the likelihood of pretransplant death for patients awaiting liver transplantation. In this study, as an alternative way to estimate this likelihood, we propose classification and regression tree techniques for censored survival data with competing risks. We develop several types of trees and use simulation to compare the results.

Extensions of Frailty Models for Multivariate Grouped Survival Data

✱Andrea B. Troxel, University of Pennsylvania, 423 Guardian Dr, 632 Blockley Hall, Philadelphia, PA 19104, *atroxel@mail.med.upenn.edu*; Denise A. Esserman, The University of North Carolina at Chapel Hill

Key Words: frailty, multivariate, random effects, quality of life

Quality of life (QOL) has become an important outcome in understanding treatment effects, especially in the palliative care setting. Thus, patients making treatment decisions should consider QOL as well survival. To facilitate this assessment, we propose a shared frailty model that jointly models grouped (QOL) and continuous survival data. We focus on the univariate normal distribution to model the random frailty. Via simulations, we explore the robustness of the model to misspecification of the frailty distribution using maximum likelihood (ML) and residual maximum likelihood (REML) methods. In addition, we explore the impact on parameter estimates of different amounts of censoring for the continuous outcomes. We present an example using a clinical trial in which both survival and QOL data were collected.

Censored Geometric Regression with Application to Human Subfertility

✱Steven J. Schwager, Cornell University, Department of Biological Statistics & Comp Bio, 1194 Comstock Hall, Ithaca, NY 14853, *sjs5@cornell.edu*; Daniel Zelterman, Yale University

Key Words: Bayesian analysis, censored discrete distribution, EM algorithm, generalized linear models, overdispersion, survival data

The geometric distribution describes the number of failures observed in an iid Bernoulli sequence until the first success. We propose a model for geometric distributed data in the framework of generalized linear models, incorporating right-censored responses and covariates to describe the parameters of the distribution. Diagnostics include a test for overdispersion. These methods are applied to a model of the success of pregnancies induced by intrauterine insemination in 200 women who were treated for subfertility at Yale. The number of services performed until success is modeled as a geometric distribution. Right censoring occurs when a woman discontinues treatment. The parameters of the geometric model and the probability of discontinuing treatment are estimated by an EM algorithm. A Bayesian model is developed to derive predictions for the number of treatments expected before success.

Models for Longitudinal Survival Data

✱Stuart R. Lipsitz, Harvard Medical School, Division of General Internal Medicine,, 1620 Tremont St. ,3rd floor, BC3 002D, Boston, MA 02120-1613, *Slipsitz@partners.org*; Debajyoti Sinha, Florida State University; M. Brent McHenry, Bristol-Myers Squibb Company

Key Words: frailty, Kendall's-tau

Longitudinal studies where the outcome at each occasion is a failure time are becoming common. For example, an AIDS clinical trial was undertaken to determine the effects of different doses of a treatment in preventing HIV-1 virus positivity over time. Blood samples were collected at weeks 4, 8, and 12 from randomization; for a given blood sample, the failure time is the days until the virus was detected in the blood sample. Frailty models are popular for modeling correlated survival data. An acknowledged limitation of a frailty model is that it only allows one random effect for the correlated survival data on each subject. Here, we propose use of separate, but correlated, frailties across occasions. The proposed model allows for a flexible correlation structure for the correlated frailties, such as autoregressive. The correlations can be interpreted in terms of Kendall's-tau.

256 Beyond Design: Reaching Out with NHANES Data ●▲

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section, WNAR, Section on Health Policy Statistics

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

When Good Weights Seem Bad: Nuances of Sample Weights in the NHANES

✱Lester R. Curtin, Centers for Disease Control and Prevention, 3311 Toledo Road, Hyattsville, MD 20782, *lrc2@cdc.gov*; Sylvia Dohrmann, Westat

Key Words: Sample Weights, Nonresponse, Target Population, Design Based Estimation

While most analytic applications using public data files are well done, there is potential for misuse and misinterpretation due to the complexity of the data base. Most large national surveys are designed to meet multiple analytic objectives yet, for ease of use and consistency of results, a single sample weight is provided to the public for design based estimation. Issues of use and interpretation of the design based estimates can arise due to misunderstanding of the target population, design efficiencies versus inefficiencies for certain characteristics, item non-response, and other factors. This presentation provides examples from the NHANES data where the sample weight is misused or misinterpreted in published articles. Examples include applications to specific population characteristics (such as income), subnational estimates, and analyses where item nonresponse is a factor.

Producing Local Area Estimates for NHANES

✱Leyla Mohadjer, Westat, 1650 Research Blvd., Rockville, MD 20850, *LeylaMohadjer@Westat.com*; Robert E. Fay, Westat; Lester R. Curtin, Centers for Disease Control and Prevention

Key Words: Small area estimation, Domain estimation

The National Health and Nutrition Examination Survey (NHANES) is one of a series of health-related surveys that have furnished a rich source of national data for over 50 years. A unique feature of NHANES is the complete medical examination carried out in Mobile Examination Centers (MECs) that travel to survey locations (mostly individual counties) in the country. Thus the survey requires extensive cooperation of sampled counties to locate and place the MECs in appropriate sites. In addition, NHANES has an extensive outreach program, including local media coverage and contacts with local organizations, to reach as many sampled persons as possible. To minimize the risk of county-level refusal, we consider the question of whether counties can be rewarded for their cooperation by providing them with selected county-level estimates. We sketch a general approach to do so, combining approaches.

Challenges and Consequences from Combining National Health Surveys with Air Monitoring Data

✱Jennifer D. Parker, Centers for Disease Control and Prevention, NCHS, 3311 Toledo Road, Hyattsville, MD 20782, *jdparker@cdc.gov*; Lester R. Curtin, Centers for Disease Control and Prevention; Nataliya Kravets, NOVA Research

Key Words: NHANES, air pollution

The objective of this paper is to describe the analytic challenges of data files obtained by linking air quality data from the EPA with health data from the NHANES. The combination of health data with routinely monitored environmental exposure data has the potential to enhance our understanding of environmental impacts on health. However, air monitors are placed for regulatory not scientific purposes and may not correspond to areas sampled in national surveys. As a result, between 20% and 70% of the NHANES respondents may be linked, depending on the pollutant, the spatial and temporal criteria for exposure assignment, and the location of the NHANES sample units. The demographic profile of the linked records differs from that of the full NHANES sample.

Comparing Methods for Left-Censored Data Using Two-Parameter Lognormal and Johnson's SB Distributions: The Case of Environmental Data from NHANES

✱Philip Villanueva, U.S. Environmental Protection Agency, 1200 Pennsylvania Avenue, NW, Office of Pesticides Programs, Mailcode 7509P, Washington, DC 20460, *villanueva.philip@epa.gov*

Key Words: maximum likelihood estimation, enviromental data, pesticides, lognormal, Johnson's SB, censored data

Concentrations of chemical compounds, such as pesticides, are found in various environmental and biological media. Some pesticide concentrations are reported below the analytical limit of detection (LOD), even though the pesticide may be present in small, but undetectable amounts. Concentrations reported below LOD are referred to as left-censored. Maximum likelihood estimation (MLE) can incorporate information regarding the proportion of the dataset that is reported below the LOD to fit various probability distributions to such left-censored data. This paper provides a comparison of fitting the lognormal and Johnson's SB distributions to pesticide concentrations obtained from NHANES biomonitoring data. The Office of Pesticide Programs will continue to explore and analyze NHANES biomonitoring data to help inform its pesticide risk assessments in the future.

257 Statistical Innovations in Health Services Research

Section on Health Policy Statistics, Section on Quality and Productivity, WNAR

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

A Stochastic Frontier Model for Cost-Effectiveness Analysis with Clustered Data

✱Zhehui Luo, RTI International, 27709, *zluo@rti.org*; Jeremy W. Bray, RTI International; Alexander Cowell, RTI International; Joseph C. Gardiner, Michigan State University

Key Words: cost-effectiveness, stochastic frontier, clustered data, group randomization, hierarchical model

A stochastic cost-effectiveness analysis (CEA) uses sampled individual-level costs and effects which allows consideration of sampling variability in the estimation of desired decision parameters. However, costs and effects are commonly estimated independently or via seemingly unrelated regressions. As a result the uncertainty in both costs and effects due to unobserved heterogeneity versus health production inefficiency at the patient level are not distinguished. We propose a stochastic frontier model that explicitly connects the estimation of costs to the production of health and allows separation of patient heterogeneity from inefficiency. We apply the method in a multi-center randomized controlled trial for treatment for alcohol problems. Our model implies an additional source of uncertainty in CEA and has consequences for decision uncertainty modeling.

Improved Health Policy Decisionmaking via Efficiency Gains with Recurrent Events

✱Edsel A. Pena, University of South Carolina, Department of Statistics, University of South Carolina, Columbia, SC 29208, *pena@stat.sc.edu*; Akim M. Adekpedjou, University of Missouri-Rolla; Jonathan Quiton, Western Kentucky University

Key Words: recurrent monitoring, Koziol-Green model, sum-quota accrual

Health policy decisions are typically based on the outcomes of statistical analysis of observed health and biomedical data. With reliable inference from such health data, then reliable policy decisions will be made. Oftentimes, in the context of failure time monitoring, statistical inference methods are usually based on the time-to-first event, and the usual methods of analysis are based on the Kaplan-Meier estimator and Cox regression. However, in many settings, the event of interest is recurrent. In this situation, we illustrate the potential gain in efficiency when doing inference by considering all the observed recurrences, and how this could translate to improved health policy decisions. We do this efficiency analysis by extending the Koziol-Green model, which was used in efficiency analysis for single-event models.

Use of Spatially Adjusted Bayesian Additive Regression Tree (SBART) Model To Reduce Ecological Fallacy in Health Services Research

✱Ya-Chen T. Shih, The University of Texas M.D. Anderson Cancer Center, Section of Health Services Research, Dept of Biostatistcs, 1515 Holcombe Blvd. #447, Houston, TX 77030, *yashih@mdanderson.org*; Song Zhang, The University of Texas Southwestern Medical Center

Key Words: disparities, ecological fallacy, Bayesian additive regression tree model

Area-based statistics (e.g., income at zip code level) are often used as proxy SES variables in research lacking person-level data on SES. This census-based approach may lead to ecological fallacy; we applied the SBART model to explore this issue. The SBART incorporates spatial random effects to BART. It has been shown that compared with the census-based approach, estimates obtained from the SBART were much closer to the true estimates (i.e., from data containing information on person-level SES). We applied the SBART method to examine the racial/ethnic disparities in survival among colorectal cancer patients using the link-SEER-Medicare data, with median household income at zip code level as proxy for SES. The comparison between SBART and census-based methods confirmed our concern of ecological fallacy in that SBART reported a weaker association between race/ethnicity and survival.

A New Approach for the Analysis of Stage-Sequential Development

✱Hwan Chung, Michigan State University, B601 West Fee Hall, East Lansing, MI 48824, *hchung@epi.msu.edu*

Key Words: stage-sequential process, latent class analysis, Longitudinal data, categorical data

In many drug-related studies, the identification of subgroups of individuals who follow similar stage-sequential development has become an issue of primary importance. This study proposes a new latent variable model, the Latent Class-Profile Analysis (LCPA), which provides a set of principles for systematic identification of homogeneous subgroups in their stage-sequential development. LCPA divides identification processes into two levels: LCPA identifies latent classes at the first level and latent class-profiles at the second level. All individuals in a second-level subgroup are expected to have the similar sequential development over time. In this study, we explore a variety of theoretical and practical issues surrounding the use of LCPA. We illustrate these issues with an example from the adolescent substance use data set.

258 Data Analysis Issues in Medical Device Trials ●▲

Biopharmaceutical Section, WNAR, Biometrics Section
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Subgroup Analysis

✱Chul H. Ahn, U.S. Food and Drug Administration, 1350 Piccard Drive, Rockville, MD 20850, *chul.ahn@fda.hhs.gov*

Key Words: data dredging, multiple testing, regression to the mean, statistical power, subgroup p-value, interaction test

A subgroup analysis is a way to identify groups of patients who respond “differently” to treatment. It appears in almost every device or drug trials. The one big question in this analysis is whether the active treatment is ineffective or worse in some subgroups while being beneficial within others. The subgroup analyses should be pre-planned, however, unfortunately, many clinical trials are not designed, and sized, with subgroup analyses in mind; these analyses are usually done after the overall analysis is completed (post hoc). In this presentation, we will consider five major problems in subgroup analysis among others—data dredging, multiple testing, regression to the mean, statistical power, and subgroup p-value. They will be presented along with examples.

A Comparison of Noninferiority Testing Methods for Survival Endpoints

✱Brian P. Johnson, Boston Scientific Corporation, *johnsob8@bsci.com*

Key Words: Survival analysis, noninferiority testing, score function, and log-rank statistic

The log-rank test has long-been a standard in the analysis of time-to-event data and hypotheses of noninferiority have become increasingly important in the relative assessment of treatment modalities. A hazard ratio that is constant over time is often the parameter of interest. Under this scenario, Com-Nougue et al. (1993) consider noninferiority testing, given censored data, in terms of a modified log-rank statistic. Nam et al. (2005) derive a score test for the same, while Jung et al. (2005) develop a log-rank statistic based upon the partial score function and provide a formula for sample size estimation. We expect the operating characteristics of the three tests to differ when the hazard ratio is constant and present empirical results when it is made to change over time. The methods are applied to typical survival data from clinical trials of drug-eluting stents (DES).

Group Sequential Trial Design To Test Both Superiority and Noninferiority

✱Vandana Mukhi, U.S. Food and Drug Administration, 1350 Piccard Dr., CDRH, Rockville, MD 20850, *vandana.mukhi@fda.hhs.gov*; Yongzhao Shao, Iowa State University

Key Words: Superiority, Non-Inferiority, Group Sequential, Phase III Trial, Design, Three-outcome

Specification of the study objective of superiority or non-inferiority at the design stage of a pivotal clinical trial can sometimes be very difficult due to the uncertainty that surrounds the efficacy level of the experimental treatment/device. This uncertainty makes it tempting for investigators to design a trial that would allow testing of both superiority and non-inferiority hypotheses simultaneously. In this talk we discuss a group sequential trial design for testing superiority and non-inferiority if it is deemed plausible. This is in joint work with Dr. Yongzhao Shao.

Application of Bootstrap and Jackknife Methods in Clinical Risk Score Generation

✱Aijun Song, Boston Scientific Corporation, 4100 Hamline Avenue North, St. Paul, MN 55112, *songs1@bsci.com*; Yun Lu, Boston Scientific Corporation; Kevin Najarian, Boston Scientific Corporation

Key Words: Bootstrap, Backward Selection, Stepwise Selection, Risk Score Generation, Jackknife

Variable selection is an important step in any predictive model. It is often conducted by backward or stepwise selection procedures. Because sampling variation is not considered, the result from one sample using either stepwise or backward selection can be very unreliable. It may lead to overly optimistic estimates with respect to predictive ability and model fit. Problems grow as the sample size becomes small. Bootstrap and Jackknife method have been recommended in combination with backward/stepwise selection procedure for the predictor selection process to measure and correct overly optimistic inferences obtained from analyzing a single dataset. This analysis will investigate and compare the performance of these two methods and the impact on the selection of predictors. It will be conducted on data from a post-marketing, drug-eluting stent registry.

Covariate-Adjusted Putative Placebo Analysis in Active-Controlled Clinical Trials

✱Zhiwei Zhang, U.S. Food and Drug Administration, 1350 Piccard Dr., CDRH, HFZ-550, Rockville, MD 20850, *zhiwei.zhang@fda.hhs.gov*

Key Words: conditional effect, constancy, causal inference, discounting, non-inferiority, potential outcome

Even though an active-controlled trial provides no information about placebo, investigators and regulators often wonder how the experimental treatment would compare to placebo should a placebo arm be included in the study. A putative placebo analysis attempts to address this question by combining information from previous studies comparing the active control

with placebo. Such an analysis often requires a constancy assumption, namely that the control effect relative to placebo is constant across studies. When the constancy assumption is in doubt, there are ad hoc methods that "discount" the historical data in a conservative fashion. This paper presents a different approach that does not require constancy or involve discounting, but rather attempts to adjust for any imbalances in covariates between the current and historical studies.

259 Statistics for Climate Models ●

Section on Bayesian Statistical Science, Section on Statistics and the Environment
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Learning About Complex Physical Systems from Multiple Computer Models

✱Leanna L. House, University of Durham, Mathematical Sciences Department, Science Labs., South Road, Durham, DH1 1RS United Kingdom, *leanna.house@durham.ac.uk*; Michael Goldstein, University of Durham; Jonathon Rougier, University of Bristol

Key Words: Computer Models, Multiple Models, Bayes Linear, second Order Exchangeable

Tough scientific questions tend to inspire the design of multiple computer models to simulate the same physical process. Thus, a set of computer models, to which we refer as a Multiple Model Ensemble (MME), often exists whose elements rely fundamentally on common scientific principles, but differ for reasons including designer assumptions, resources, and complexity. Because of the differences, typical uncertainty analyzes of deterministic models tend to rely solely on the output from one model (the model in question) and exclude the results pooled from an entire ensemble. In doing so, valuable information pertaining to the expected value and variance for the discrepancy between a chosen computer model and reality is overlooked. Our goal is to develop a flexible framework for using an MME sensibly to learn about the uncertainty of a chosen model.

Computional Approaches for Parameter Estimation in Climate Models

Alejandro Villagran, University of New Mexico; ✱Gabriel Huerta, University of New Mexico, 87131, *ghuerta@stat.unm.edu*; Charles Jackson, University of Texas at Austin; Mrinal Sen, University of Texas at Austin

Key Words: Geophysical inverse problem, Simlated Annealing, Bayesian Inference, Climate models, Adaptive methods

To quantify the uncertainties arising in climate prediction it is necessary to estimate a multidimensional probability distribution. The computational cost of evaluating such probability distribution is impractical using traditional methods (Metropolis/Gibbs algorithms). Multiple Very Fast Simulated Annealing (MVFSA) is a fast but approximate algorithm that has been used by geophysicists for highly nonlinear problems in geophysics. The purpose of this talk is to describe the development and application of different computational techniques used in geophysical inverse problems. Besides MVFSA, we apply Adaptive Metropolis methods to a surrogate climate model that is able to approximate the noise and response behavior of a realistic atmospheric general circulation model. We show that adaptive methods are superior to MVFSA to approximate the posterior distribution of a climate model.

Hierarchical Calibration of a Climate Model

✱James Gattiker, Los Alamos National Laboratory, PO Box 1663, Los Alamo, NM 87544, *gatt@lanl.gov*; Dave Higdon, Los Alamos National Laboratory

Key Words: calibration, simulation analysis, climate modeling, bayesian methods

This paper presents an extended framework for analysis of systems where simulation model experiments correspond to observations of real system experiments. We extend the Bayesian Gaussian process emulator methodology to calibration in the case where multiple types of observations or models (i.e., model diagnostics) are available. The challenge in multidiagnostic calibration is trading off the areas of consistency and areas of conflict or tension in the calibrations from the diagnostics. The solution we present is a hierarchical model on the calibration parameters, and the corresponding Bayesian framework for analysis. This method demonstrated with calibration of an Atmosphere-Ocean General Circulation Model dataset, calibrating parameters including climate sensitivity. Three data set diagnostics are derived from the simulation model and calibrated to corresponding observation data sets.

Estimating the Probability Distribution of Climate System Properties

✱Bruno Sanso, University of California, Santa Cruz, University of California, 1156 High Street, Mail Stop SOE2, Santa Cruz, CA 95064, *bruno@ams.ucsc.edu*; Chris Forest, Massachusetts Institute of Technology

Key Words: climate model, computer models, bayesian methods

The behavior of modern climate system simulators is controlled by numerous parameters. By matching model outputs with observed data we can perform inference on such parameters. This is a calibration problem that usually requires the ability to evaluate the computer code at any given configuration of the parameters. As the climate system simulator attempts to describe very complex physical phenomena, the task of running the model is very computationally demanding. Thus, a statistical model is required to approximate the model output. In this work, we use output from the MIT 2D climate model (MIT2DCM), historical records and output from a three-dimensional atmospheric ocean general circulation model (AOGCM), to obtain estimates of the effective thermal diffusivity in the deep-ocean, the climate sensitivity and the net aerosol forcing that control the MIT2DCM behavior.

Parameter Estimation for Computationally Intensive Nonlinear Regression with an Application to Climate Modeling

✱Dorin Drignei, Oakland University, MI 48309, *drignei@oakland.edu*; Doug Nychka, National Center of Atmospheric Research; Chris Forest, Massachusetts Institute of Technology

Key Words: Equilibrium climate sensitivity, Observed and modeled climate, Space-time modeling, Statistical surrogate, Temperature data

Nonlinear regression is a useful statistical tool, relating observed data and a nonlinear function of unknown parameters. When the parameter-dependent nonlinear function is computationally intensive a straightforward regression analysis by maximum likelihood is not feasible. The method presented in this paper proposes to construct a faster running surrogate for such a computationally intensive nonlinear function, and to use it in a related nonlinear statistical model that accounts for the uncertainty associated with this surrogate. A pivotal quantity in the Earth's climate system is the climate sensitivity: the change in global temperature due to doubling of atmospheric CO2 concentrations. This, along with other climate parameters, are estimated by applying the statistical method developed in this paper, where the computationally intensive nonlinear function is the MIT 2D climate model.

260 Dimension Reduction, Variable Selection, and Correlation Pursuit in Semiparametric Settings ●▲

Section on Nonparametric Statistics, WNAR, IMS
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Dimension Reduction with Possible Application in Bioinformatics

✱Wenxuan Zhong, University of Illinois at Urbana-Champaign, 119 Illini Hall, 725 South Wright Street, Champaign, IL 61822, *wenxuan@uiuc.edu*

Key Words: Correlation Pursuit, Motif Finding, Variable Selection, Dimension Reduction

Information for regulating a gene's transcription is contained in the conserved patterns (motifs) on the upstream/downstream DNA sequence (promoter region) close to the target gene. By combining the information contained in both gene expression measurements and genes' promoter sequences, I proposed a novel procedure for identifying functional active motifs under certain stimuli. A dimensional reduction model was used to associate promoter sequence information of a gene and its mRNA expression measurements. A correlation pursuit (corps) variable selection method was developed to select the potential functional active promoter sequences. In this talk, I will demonstrate the advantage of the corps both theoretically and empirically.

A Link-Free Method for Testing the Significance of Predictors

✱Peng Zeng, Auburn University, Department of Mathematics & Statistics, Auburn University, Auburn, AL 36849, *zengpen@auburn.edu*

Key Words: dimension reduction, variable selection, central subspace

One important step in regression analysis is to identify significant predictors from a pool of candidates so that a parsimonious model can be obtained using these significant predictors only. However, most of the available methods assume linear relationships between response and predictors, which may be inappropriate in some applications. In this talk, we discuss a link-free method that avoids specifying how the response depends on the predictors. Therefore, this method has no problem of model misspecification, and it is suitable for selecting significant predictors at the preliminary stage of data analysis. A test statistic is suggested and its asymptotic distribution is derived. Examples are used to demonstrate the proposed method.

An Asymptotic Analysis of the Stepwise Correlation Pursuit Variable Selection Method

✱Tingting Zhang, Harvard University, 1 Oxford Street 608, Cambridge, MA 02138, *tzhang@fas.harvard.edu*; Wenxuan Zhong, University of Illinois at Urbana-Champaign; Jun S. Liu, Harvard University

Key Words: variable selection, stepwise regression, single index model

Classic nonparametric methods for regression analysis break down quickly as the dimension of predictors increase. As an alternative approach, a stepwise correlation pursuit variable selection procedure for single index model has been proposed recently by Zhong et. al. This article aims to study the asymptotic behavior of this stepwise variable selection procedure. More specifically, we have analyzed the convergence rate of the test statistics under the null hypothesis of no effect for selected predictors and the power of each testing step as sample size goes to infinite. We have also compared the new variable selection procedure with the classic stepwise forward addition backward deletion procedure for linear regression models. The possibility of generalizing the method to multiple index model is also discussed in the paper.

Selective Partially Augmented Naïve Bayes with Model Averaging

✱Yuan Yuan, Harvard University, MA 02138, *yyuan@stat.harvard.edu*; Jun S. Liu, Harvard University

Key Words: Naive Bayes, Bayes network, Variable selection

Naïve Bayes is a classifier that assumes conditional independence of all features (covariates) given the class label. Although this assumption is unrealistic, the naïve Bayes classifier is competitive with some state-of-the-art classifiers. Many methods have been developed to relax this independence assumption to achieve better classification accuracy. Here we propose a Selective Partially Augmented Naïve Bayes model (SPAN) which considers the joint distribution of small groups of selected features in an augmented network structure. Feature and network structure selection is achieved by applying Markov chain Monte Carlo methods. Instead of inferring one single model, Bayesian model averaging is used to make prediction (classification). Our model is especially useful in problems with large number of covariates such as DNA marker association studies.

Penalized Linear Methods for Estimation and Variable Selection in Index Models

✱Michael Y. Zhu, Purdue University, 250 N. University Street, West Lafayette, IN 47907, *yuzhu@stat.purdue.edu*; Jian Zhang, Purdue University

Key Words: index models, lasso, ridge regression, variable selection

Index models are among the most popular semi-parametric models in econometrics and statistics, mainly due to their flexibility in handling the unknown link function and resistance to the curse of dimensionality. In this work, we study the penalized linear methods for index estimation and variable selection, in particular the least squares methods with L1 and L2 penalties. These two methods are known as Lasso and ridge regression. We will show that the penalized linear methods can successfully estimate the index parameters and select the relevant variables. Theoretical results and real and simulation examples will be reported. The developed methods can be used in genetic or genomical applications where high dimensional and nonlinear data are prevalent.

261 Statistical Graphics for Analysis and Reporting of Clinical Data ●▲

Section on Statistical Graphics, Biopharmaceutical Section
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Application of Statistical Graphics in Clinical Trial Data Analysis

✱Julia Wang, Johnson & Johnson Pharmaceutical R&D, LLC, 57 Sentinel Dr., Basking Ridge, NJ 07920, *jwang@prdus.jnj.com*; Surya Mohanty, Johnson & Johnson Pharmaceutical R&D, LLC

Key Words: Statistical graphics, Splus

Well-structured statistical graphics is increasingly being used to extract and communicate information from the clinical trial data in a lucid and succinct manner. This presentation will demonstrate several effective applications of statistical graphics in clinical trial data analysis.

Multivariate Display of Laboratory Data for Safety Analysis

✱Charles F. Contant, Pfizer Global Research and Development, 50 Pequot Ave, New London, CT 06320, *charles.contant@pfizer.com*; Robb J. Muirhead, Pfizer Global Research and Development

Key Words: multivariate graphics, laboratory data, wavelet transforms, multivariate distribution

Evaluation of safety is one of the core activities of drug development. Adverse events and clinical pathology are the primary safety data types. FDA guidance indicates that descriptive statistics are the primary means of examining summaries of these data. Inspection of the raw data is used to monitor safety as a study progresses. The laboratory data are usually examined univariately and flagged when a single value exceeds some arbitrary threshold. We will examine several methods of multivariate display of laboratory data. Multivariate presentation of individual data vectors(bars, star plots etc) and methods based on grouped data(biplots and combination of basis functions) will be presented. The ability to detect abnormal values and to determine the source of the abnormality will be compared qualitively among the methods. We will make recommendations for future development.

Graphical Exploration for Cardiovascular Safety Data in Clinical Trials

✱Ihab G. Girgis, Johnson & Johnson Pharmaceutical R&D, LLC, Advanced Modeling & Simulation, 920 Route 202 South,PRD Bldg. Rm 2722, Raritan, NJ 08869, *igirgis@prdus.jnj.com*; Surya Mohanty, Johnson & Johnson Pharmaceutical R&D, LLC

Detecting drug-induced effect on cardiac repolarization or QT interval is a closely monitored safety element in drug development and more recently, it is thoroughly scrutinized in regulatory submissions. This work focuses on the graphical representation of the exploration and the modeling of baseline QT clinical data.

Reporting of Safety and Efficacy Data Collected in Clinical Trials: A Statistical Reviewer's Perspective

✱Mat Soukup, U.S. Food and Drug Administration, *Mat.Soukup@fda.hhs.gov*

Key Words: graphics, clinical trial

Increasing concerns for drug safety and the lengthy drug development process have highlighted the need for efficient, thorough and transparent statistical data analysis and reporting of clinical studies. The use of compelling statistical graphics can provide transparent representations of the data collected in clinical trials. Such presentations improve the quality of clinical decisions, increase the likelihood of discerning safety and efficacy signals, and enable more productive interactions between the U.S. FDA and regulated companies. In this presentation, we will discuss graphical displays that improve upon traditional methods for the tabular reporting of clinical trial results, as well as examples that have been critical in making regulatory decisions.

Graphical Analysis of Adverse Event Data

✱Haijun Ma, Amgen, Inc, Amgen One Ct. Dr., Thousand Oaks, CA 91320, *hma@amgen.com*; H. Amy Xia, Amgen, Inc

Key Words: safety analysis, adverse events, graphics, clinical trial safety, safety signal detection

Adverse event data play a crucial role in the analysis of clinical trial safety data. In this talk we share our experience in monitoring and analyzing clinical trial adverse event data. We examine various metrics commonly used for adverse event data analysis, approaches that are helpful for identification of adverse drug reactions, methods focused on analysis of Events of Interests and ways to incorporate adverse event data with other safety data. We also discuss challenges in implementing routine monitoring while trials are still ongoing. Statistical graphics, as a powerful statistical analysis tool, are applied as appropriate and discussed through out the talk.

262 Joint Student Paper Competition Presentations

Social Statistics Section, Section on Government Statistics, Section on Survey Research Methods

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Local Post-Stratification and Diagnostics in Dual System Accuracy and Coverage Evaluation for U.S. Census

✱Cheng Yong Tang, Iowa State University, Department of Statistics, Ames, IA 50011, *yongtang@iastate.edu*; Song Xi Chen, Iowa State University

Key Words: Capture-recapture, Census undercounts, Discrete kernel smoothing, Erroneous enumeration, Imputation, Missing Values

This paper proposes a local post-stratification approach to the dual system estimation for evaluating coverage in a Census. The local stratification is achieved by nonparametric estimators for the Census enumeration and correct enumeration probabilities, which may be employed as an alternative to the existing post-stratification in the Census dual system estimation. We propose an imputation based estimator for general nonparametric regression with missing values, which can handle both continuous and categorical covariates and is used for model diagnostic and checking. Our theoretical analysis and simulation studies indicate attractive properties of the imputation based estimator. The proposed approach is applied to analyze research data files from the 2000 Census dual system surveys.

The Prior Selection and Approximations in a Hierarchical Bayes Approach: An Application to the Small-Area Income and Poverty Estimation

✱Santanu Pramanik, University of Maryland, College Park, JPSM, 1218J Lefrak Hall, College Park, MD 20742, *spramanik@survey.umd.edu*

Key Words: Adjustment for density maximization, Fay-Herriot model, Frequentist properties, Laplace approximation, Mean likelihood, Residual maximum likelihood

We consider two important issues in the hierarchical Bayesian inference, namely the selection of prior for the variance component and highly accurate simple approximations to various posterior moments involved. We propose a prior distribution for the variance component that yields a positive estimator having good frequentist properties when posterior mode is considered as an estimator. We discuss several approximate hierarchical Bayes methods which have several advantages over the complicated Monte Carlo Markov Chain methods. By applying our methodology on data related to the SAIPE project of the U.S. Census Bureau, we demonstrate the utility and accuracy of our approach in solving real life small area estimation problem relative to some other existing methods.

Bayesian Penalized Spline Model-Based Estimation for Finite Population Proportion in Unequal Probability Sampling

✱Qixuan Chen, The University of Michigan, Department of Biostatistics, M4024 SPH II, School of Public Health, 1420 Washington Heights, Ann Arbor, MI 48109-2029 USA, *qixuan@umich.edu*; Michael R. Elliott, The

University of Michigan; Roderick J. Little, The University of Michigan

Key Words: Bayesian analysis, binary data, penalized spline regression, probability proportional to size, survey samples

We propose a Bayesian penalized spline prediction (BPSP) estimator for a finite population proportion in an unequal probability sampling setting. This new method allows the probabilities of inclusion to be directly incorporated into the estimation of a population proportion, using a probit regression of the binary outcome on the penalized spline of the inclusion probabilities. The posterior predictive distribution of the population proportion is obtained using Gibbs sampling. Simulation studies and a real example in tax auditing show that the BPSP estimator is more efficient, and its 95% credible interval provides better coverage with shorter average width than the Horvitz-Thompson (HT) and generalized regression (GR) estimators. Compared to parametric model-based prediction estimators, the BPSP estimators achieve robustness to model misspecification and influential points in the sample.

Model Selection with Partially Synthetic Data

✱Satkartar Kinney, National Institute of Statistical Sciences, *saki@niss.org*; Jerome Reiter, Duke University; James Berger, Duke University/SAMSI

Key Words: Bayesian model uncertainty, Bayes factor approximation, Data confidentiality, Multiple imputation

Several statistical agencies use, or are considering the use of, multiple imputation to limit the risk of disclosing respondents' identities or sensitive attributes in partially synthetic public use data files. Methods for obtaining inferences for scalar and multivariate estimates have been developed for partially synthetic datasets, as well as several other types of multiply-imputed datasets. Methods for conducting model selection with multiply-imputed data are very limited, and this paper moves toward filling this need. As a first step, a simple case is considered in the context of partially synthetic data where the analyst has some knowledge about the imputation procedure. In this scenario, a Bayesian approach to implementing model selection with partially synthetic datasets and a Bayes factor approximation similar to the BIC are derived and illustrated.

Calculating Cell Bounds in Contingency Tables Based on Conditional Frequencies

✱Byran J. Smucker, The Pennsylvania State University, 325 Thomas Building, University Park, PA 16802, *bjs379@psu.edu*; Aleksandra B. Slavkovic, The Pennsylvania State University

Key Words: statistical disclosure limitation, contingency tables, confidentiality, linear programming, integer programming

In the statistical disclosure limitation literature work has been done on summary and release of contingency table data in terms of evaluating the bounds on cell entries in k-way contingency tables given sets of marginal totals. Narrow bounds, especially for cells with low counts, could pose a privacy risk. The focus of this paper is measuring disclosure risk when tables of rates, that is tables with conditional probabilities, are released. We derive the closed-form solutions for the linear relaxation bounds on cell counts given full and partial conditional probabilities. We also compute the corresponding sharp integer bounds via integer programming and show that there can be large differences in the width of these bounds, suggesting that the use of linear relaxation is often an unacceptable shortcut to estimating the sharp bounds and disclosure risk for the tables of rates.

263 Research Using the Survey of Attitudes Toward Statistics (SATS)

Section on Statistical Education

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Research Using the Survey of Attitudes Toward Statistics (SATS)

✱Marjorie E. Bond, Monmouth College, Dept. of Mathematics & CS, 700 East Broadway, Monmouth, IL 61462, *mebond@monm.edu*; ✱Candace Schau, CS Consultants, 12812 Hugh Graham Road NE, Albuquerque, NM 87111, *cschau@comcast.net*; ✱Rebecca L. Pierce, Ball State University, Dept of Mathematical Sciences, Robert Bell Bldg, 465, Muncie, IN 47304, *rpierce@bsu.edu*; ✱Sue Schou, Idaho State University, PO Box 4043, Pocatello, ID 83205, *schosue@isu.edu*

Key Words: SATS, Statistics Education, Attitudes, on-line courses, survey research problems, Survey of Attitudes Toward Statistics

We will begin with a brief summary of the SATS. Panel members then will present a summary of their research using the SATS. These summaries will include comparing SATS results among students taking an on-line course versus a face-to-face course; comparing SATS results among three different majors (education, computer science, and construction management); and investigating the relationships among the six SATS attitudinal component scores and the additional items assessing global attitudes and other student information. We will discuss problems commonly found in research using the SATS and present how audience members can easily collect SATS data from students in their classes.

264 Problems with RDD Surveys

Section on Survey Research Methods, Section on Government Statistics

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Statistical Challenges Facing Cell Phone Surveys

✱Michael P. Battaglia, Abt Associates, Inc., 55 Wheeler Street, Arlington, MA 02138, *mike_battaglia@abtassoc.com*; Martin R. Frankel, Baruch College and Abt Associates Inc.; Ali H. Mokdad, Centers for Disease Control and Prevention

Key Words: Random-digit-dialing, cellular-only households

Landline random-digit-dialing (RDD) survey coverage problems due to cell phone-only households have led to a growing number of cellular telephone surveys being conducted at the national and state level to supplement traditional RDD surveys. We discuss the trade-offs of screening for cell phone-only household/adults versus using a dual frame sample design that also includes households/adults with landline and cellular telephone service. We then discuss how the National Health Interview Survey and other surveys such as the CPS and ACS can be used to develop estimates of the size of the cell phone-only household/adult population for states and other subnational geographic areas.

Adjustment for Noncoverage of Nonlandline Telephone Households in an RDD Survey

Sadeq R. Chowdhury, National Opinion Research Center; ✱Robert Montgomery, National Opinion Research Center, 55 East Monroe, Suite 2000, Chicago, IL 60603, *Montgomery-Robert@NORC.org*; Philip J. Smith, Centers for Disease Control and Prevention

Key Words: noncoverage adjustment, poststratification, RDD survey

The National Immunization Survey (NIS)—a nationwide, list-assisted RDD survey conducted by the NORC for the Centers for Disease Control and Prevention—monitors vaccination rates of children between the ages of 19 and 35 months. As in any RDD survey, non-landline telephone (i.e., no-phone and cell-only) households are not sampled in the NIS. To compensate for this noncoverage, a special adjustment (Keeter, 1995) is applied, in which households with an interruption in landline service are used to represent nonlandline households. With the increasing proportion of cell-only households among the nonlandline households, this adjustment may no longer be very effective. Also, the adjustment introduces a considerable variation in sampling weights. The paper evaluates the impact of this adjustment on bias and variance to assess the effectiveness and to identify an optimal level of adjustment.

Cell Phone–Only Household in a National Mail Survey: Who Are They?

✱Daifeng Han, Westat, 1650 Research Blvd., Rockville, MD 20850, *DaifengHan@Westat.com*; David Cantor, Westat

Key Words: Dual frame surveys, cell-only population, mail mode

In RDD surveys, how the noncoverage of cell-only population may affect survey results depends on the magnitude of the difference between cell-only respondents and those with landlines. Previous studies have shown that the cell-only population is different from that with landline telephones in terms of age, race-ethnicity, marital status and tenure. Using a dual-frame survey of mail and RDD samples, we compare the cell-only respondents and the respondents with a landline in the mail sample on a range of questions covering demographic characteristics, health characteristics, and behavior patterns on media consumption. Results are also compared to external benchmark of a field survey with higher response rate. Such comparisons allow us to assess the degree to which a mail survey can be used to generate estimates for the cell-only population.

Considerations for the Implementation of an RDD Telephone Survey

✱David Ferraro, Westat, 1650 Research Blvd., Rockville, MD 20850, *DavidFerraro@Westat.com*; Tom Krenzke, Westat; Jill Montaquila, Westat

It is well documented that nonresponse and noncoverage rates have increased steadily in random digit dial (RDD) surveys. As these rates continue to rise, there are increasing challenges to reducing bias and increasing operational efficiency. This paper discusses several methods to consider when conducting RDD surveys. We also discuss the benefits and shortcomings of the approaches—shortcomings if it is decided not to implement a particular procedure. For example, there is a growing concern with nonresponse bias; what is the impact of by-passing the household nonresponse adjustment? We also discuss the impact of not addressing cell phone only households, reducing the level of effort, and the impact other methods such as purging nonresidential numbers, advance mailings, subsampling of cases prior to refusal conversion, and other uses of auxiliary data.

Differential Bias Between the Cell Telephone–Only Population and the Landline Population in an Immunization Survey

✱Martin Barron, National Opinion Research Center, *barron-martin@norc.org*; Kirk Wolter, National Opinion Research Center; Cindy Howes, National Opinion Research Center; Karen Wooten, National Center for Immunization and Respiratory Diseases; Meena Khare, National Center for Health Statistics

Key Words: National Immunization Survey, Cell-Telephone Survey, Survey Bias

The size of the cell-telephone only population has increased in recent years raising concerns about the accuracy of RDD-based telephone surveys. Previous research on the impact of cell-telephone only on population estimates has found mixed results. In 2007, the National Immunization Survey (NIS) conducted a pilot study to determine the impact of cell-telephone-only households on childhood vaccination coverage estimates. The NIS is a nationwide, list-assisted random digit-dial (RDD) survey to monitor vaccination coverage among children aged 19 to 35 months. 49,300 cell telephone numbers in Illinois were sampled for this pilot. This paper will compare vaccination coverage estimates of children in landline households, cell-telephone only households, and combined. We discuss the implications of these results for RDD surveys, and specifically for the NIS.

Responsive Design for Random Digit Dial Surveys Using Auxiliary Survey Process Data and Contextual Data

✱Sunghee Lee, University of California, Los Angeles, 10960 Wilshire Blvd. Suite 1550, Los Angeles, CA 90024, *slee9@ucla.edu*; Tom Belin, University of California, Los Angeles; Christopher Pankonin, University of California, Los Angeles; J. Michael Brick, Westat; Jennifer Edgar, Bureau of Labor Statistics

Key Words: RDD survey, Response rate

This study shows a tailored RDD survey design by using two types of data available regardless of response status. The first coming from survey operations exhibit the history of all calls made to each sample and the indicators of survey design features, such as incentive provision. The second is prepared by linking geographic identifiers of all sample cases to the external data. As regarding survey response behaviors as a stochastic process, we fit models to predict how response behaviors change given hypothetical calling schedules and design features. The model is applied to any sample with the same set of variables, and the response behavior of a new sample can be predicted before fielding the survey. Based on the expected response behavior, the design features and calling schedules may be tailored for each case so as to maximize positive response behavior for fixed costs.

Where Have All the Smokers Gone? A Simple Strategy for Targeting Rare Subgroups for RDD Studies

✱Mansour Fahimi, Marketing Systems Group, 565 Virginia Drive, Fort Washington, PA 19034-2706, *mfahimi@m-s-g.com*; Douglas Currivan, RTI International; Jeniffer I. Perez, RTI International; Matthew Farrelly, RTI International

Key Words: RDD, CHAID, Oversampling, Design Effect

RDD studies are facing growing challenges, including eroding rates of coverage and response. Moreover, when the study population includes rare subgroups, surveys can require significant resources for screening. It is estimated that about 18% of New York adults are current smokers, where per CDC a current smoker has smoked at least 100 cigarettes and currently smokes. In order to increase the rate of locating smokers in New York, a segmentation technique was applied to prior survey data to identify high smoking areas. Subsequently, telephone exchanges associated with such areas were targeted

to increase the number of interviews with smokers. The employed disproportionate stratified sample design was implemented while managing the anticipated variance inflation due to unequal selection probabilities. This paper provides an overview of this design along with a summary of key findings.

265 Imputation for Item Nonresponse ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Imputation of Missing Data for the Pre-Elementary Education Longitudinal Survey

✱Lin Li, Westat, 1650 Research Blvd., Rockville, MD 20850, *LinLi@Westat.com*; Annie Lo, Westat; Greg Norman, Westat; Hyunshik Lee, Westat

Key Words: Complex questionnaire design, Skip patterns, Correlation structure, Complex missing patterns

In the Pre-Elementary Education Longitudinal Survey (PEELS), imputation of item missing data was done using AutoImpute (AI) software, which uses semiparametric modeling to form imputation classes. In this paper we summarize PEELS experience with AI, investigate the bias aspect of the imputed data for the PEELS teacher questionnaire data, and study the variance estimation of imputed data using multiple imputation by AI. In the study of variance estimation, we look into the bias issue for the multiple imputation method and also the performance of AI multiple imputation on domain estimation.

A Study of Imputation Alternatives for the Quarterly Financial Report

✱Melvin J. McCullough, U.S. Census Bureau, 4600 Silver Hill Road, 6K154B, MD 20746, *Melvin.McCullough@Census.Gov*; Donald M. Luery, U.S. Census Bureau; Terry L. Pennington, U.S. Census Bureau

Key Words: imputation, bias, standard error

The Quarterly Financial Report collects income and balance sheet data for most manufacturing corporations and for large mining, wholesale trade, and retail trade corporations. Historically, imputation for nonrespondent certainty cases consisted of carrying forward prior quarter data, and adjustments were made for non-respondent noncertainty cases by reweighting respondent data. Two problems with these historical methods are: a failure to adequately account for market changes, and inconsistent adjustment across companies for unit non-response. We conducted a simulation to test an alternative imputation method that uses cell means and ratios of current to prior quarter respondent data. This paper describes the design and implementation of this simulation and presents an analysis of the results. The proposed method yields estimates with smaller bias and comparable standard errors.

Fractional Regression Nearest Neighbor Imputation

✱Minhui Paik, Iowa State University, *minhui@iastate.edu*

Key Words: missing data, imputation, replicate variance estimation

Surveys frequently have missing values for some variables for some units. Imputation is a widely used method in sample surveys as a method of handling missing data problem. We provide an new imputation procedure for various imputation models retaining many of the desirable properties of model-based imputation estimation and hot-deck imputation under fractional imputation. The main objective of this procedure is to construct an easy-to-use data set for general purpose. When the estimated imputation model is a simple linear model, the imputed estimator constructed by the proposed procedure is called the Fractional Regression Nearest Neighbor Imputation (FRNNI) estimator. In addition, a computationally efficient variance estimator is given to provide the construction of general purpose replicates for variance estimation.

Study of Item Nonresponse for a Sample of Questionnaires for the 2007 Commodity Flow Survey (CFS)

✱Hossain Eftekhari-Sanjani, U.S. Department of Transportation, Research and Innovative Technology Administration, 1200 New Jersey Ave. SE, Room E34-462, Washington, DC 20590, *hossain.sanjani@dot.gov*; Ronald Duych, U.S. Department of Transportation; Joy Sharp, U.S. Department of Transportation; Michael Margreta, U.S. Department of Transportation

Key Words: Item Non-Response, Establishment Questionnaire, Commodity Flow Survey, Transportation Survey, Establishment Survey, Data Quality

The CFS is a quinquennial, nationwide survey that collects data on the shipment of goods within the United States. It is a component of the Economic Census, and is conducted jointly by the U.S. Department of Transportation's Research and Innovative Technology Administration, Bureau of Transportation Statistics, and the U.S. Department of Commerce's Census Bureau. A sample of the 2007 CFS completed questionnaires were selected for in-depth investigation of missing or obviously bad responses to items. The CFS questionnaire is complex, and prior to implementation of the 2007 survey, extensive efforts were undertaken to improve comprehension of the questionnaire. Collected data will be checked to investigate correlates between and among missing items.

Data-Driven Post Modeling Sensitivity Analysis To Detect Missing Data Mechanism

✱Mortaza Jamshidian, California State University, Fullerton, Department of Mathematics, 800 N. State College Blvd., Fullerton, CA 92834, *mori@fullerton.edu*

Key Words: Confidence interval, Generating missing, Incomplete data, Missing at random, Missing not at random, Structural equations

Ignoring the missing data mechanism, especially when data are missing not at random (MNAR), can result in biased inference. Current methods use sensitivity analysis, via modelling the response mechanism, to assess the missing data mechanism. We proposes a method that does not require modeling the response mechanism, but rather compares parameter estimates of a measurement model fit to all of the data and to a selected subset of the data. Discrepancy in the parameter estimates of the two sets is assessed via confidence intervals, and is utilized as a tool to detect data that are not MCAR. Simulation studies and an application of the proposed method to a real data set to assess the methodology will be presented.

The Fraction of Missing Information as a Tool for Monitoring the Quality of Survey Data

✱James Wagner, The University of Michigan, G394 Perry, PO Box 1248, Ann Arbor, MI 48106-1248, *jameswag@isr.umich.edu*

Key Words: Sample Surveys, Nonresponse, Missing Data, Data Quality

Surveys have focused on the response rate as a key measure of data quality. As a result, survey research organizations have sought to maximize the response rate. However, nonresponse bias is a function of not only response rates, but also the differences between responders and nonresponders. New measures of survey data quality are needed. The fraction of missing information is one possible metric. Developed along with methods for analysis with missing data, this measure uses a model to quantify our uncertainty about the values we would impute for nonresponders. The fraction of missing information goes beyond the response rate by also involving the survey data. Therefore, this measure may be more useful for guiding survey research organizations toward collecting the highest quality data. Data from the National Survey of Family Growth are used to demonstrate the use of this metric.

General Methods and Algorithms for Modeling and Imputing Discrete Data Under a Variety of Constraints

✱William E. Winkler, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233-9100, *william.e.winkler@census.gov*

Key Words: data quality, loglinear model fit, missing data, convex constraints

Loglinear models for missing discrete data involve minor extensions of hot-deck methods (Little and Rubin 2002). Edits are structural zeros that forbid certain patterns (Fellegi and Holt 1976, Winkler 2003). This paper provides a general framework, methods, and algorithms for modeling/edit/imputation under convex constraints. The methods can be used for statistical matching (D'Orazio, Di Zio, and Scanu 2006), edit/imputation in which models are also controlled for external constraints such as in benchmark data and partially control for non-ignorable nonresponse in many situations, and for creating synthetic data with significantly reduced risk of re-identification (Winkler 2007). The software can easily be applied, uses far less resources in building productions systems, allows variance estimation and evaluation, and consistently provides better survey results (joint distributions).

266 Missing Data Methods

Biometrics Section, Section on Survey Research Methods

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Using an Informative Missing Data Model To Predict the Ability To Assess Recovery of Balance Control After Spaceflight

✱Alan H. Feiveson, NASA, Johnson Space Center, Mail Code SK3, Houston, TX 77058, *alan.h.feiveson@nasa.gov*; Scott Wood, Universities Space Research Association; William H. Paloski, NASA; Varsha Jain, Imperial College, London

Key Words: ROC, beta distribution, missing data, latent variable, random effects, balance control

To assess transient balance-control disturbances in astronauts after flight, their ability to maintain a stable upright stance was measured under several challenging stimuli using a moving-base platform. Performance was quantified by "equilibrium" scores (EQs), transformed maximum postural sway angles scaled from 0 to 100, where 100 represents perfect control (no sway) and 0 represents loss of balance (no meaningful sway angle is observed). By comparing post- to pre-flight EQs for astronauts vs. controls, we built a classifier to decide whether an astronaut has recovered. Future diagnostic performance depends on both the sampling distribution of the classifier and the distribution of its input data. Taking this into consideration, we constructed a predictive ROC by simulation, after modeling P(EQ = 0) in terms of a latent EQ-like beta-distributed random variable with random effects.

Pattern Mixture Models Incorporating Reasons for Dropout

✱Sachiko Miyahara, University of Pittsburgh, 127 Parran Hall, 130 DeSoto Street, Pittsburgh, PA 15261, *sam976@pitt.edu*; Meredith Lotz, University of Pittsburgh; Gong Tang, University of Pittsburgh

Key Words: Missing Data, Pattern Mixture Model, Nonrandom Dropout

Pattern mixture models (PMM) stratify the dataset by dropout patterns and model the repeated measures within each pattern. We propose a hierarchical PMM (HPMM) that incorporates not only dropout patterns but also various dropout reasons. Participants are classified into random and nonrandom dropout groups and the dataset is stratified according to dropout patterns within each group. Since the completers' potential dropout reasons are unobserved, an ad-hoc latent variable model is proposed to classify them into the dropout groups. Using an example dataset, estimates from this ad-hoc method are compared to MLEs for the HPMM. These methods are subsequently compared to the standard PMM and a selection model assuming ignorable dropouts. Since classification of dropout reason is subjective, a simulation study is also conducted to examine the potential bias from group misclassification.

A Unified Capture-Recapture Model

✱Matthew Schofield, Columbia University, Room 1005, MC 4690, 1255 Amsterdam Avenue, New York, NY 10027, *ms3727@columbia.edu*; Richard Barker, University of Otago

Key Words: Capture-recapture, Bayesian

Capture-recapture models have long been used to give biologists information about population dynamics. In the past 10–20 years there has been a proliferation of capture-recapture methods based primarily around increasingly complex data. We show how thinking of the capture-recapture experiment in terms of a missing data problem can (i) help to unify the abundance of models into a common framework and (ii) makes estimation of models that describe complex population dynamics more accessible to biologists.

Assessing Bias Associated with Missing Data from Joint Canada/U.S. Survey of Health: An Application

✱Sunita Ghosh, Alberta Cancer Board, 104-2711 112th St., Edmonton, AB T6J 4M1 Canada, *sunita.ghosh@usask.ca*; Punam Pahwa, University of Saskatchewan

Key Words: Missing data, Multiple Imputation, JCUSH, Survey, nonresponse

Missing data are common in large-scale surveys, arising mainly due to nonresponse in cross-sectional studies. Results are biased when missing data are ignored at the analysis stage. The present study aims to determine the bias associated when missingness is ignored and to investigate whether or not multiple imputation technique is a possible solution to address the issue of bias. The objective of the study will be achieved using the Public Use Micro Data File of Joint Canada/U.S. Survey of Health (JCUSH). JCUSH is a cross-sectional survey, which started collecting data in November 2002 and ended March 2003. The final sample contains 8,866 participants: 3,505 Canadian and 5,183 American participants. The bias will be tested using (i) available cases (only complete information), (ii) complete cases only (removing incomplete information), and (iii) multiple imputation method.

Modeling of Longitudinal Biomarker Data with Dropout and Death Using a Weighted Pseudo–Maximum Likelihood Method

✱M. Abdus Sattar, University of Pittsburgh, 305 S Fairmount St,, Apt 1, Pittsburgh, PA 15232, *upsattar@yahoo.com*; Lisa Weissfeld, University of Pittsburgh; Gina D'Angelo, Washington University School of Medicine

Key Words: Pseudo-likelihood, longitudinal data, nonignorable missingness, logistic regression

In a longitudinal study of biomarker data collected during a hospital stay, observations may be missing due to administrative reasons, the death of the subject, or the subject's discharge from the hospital, resulting in nonignorable missing data. Standard likelihood-based methods for the analysis of longitudinal data (e.g., mixed model) do not have a mechanism to account for the different reasons for missingness. Rather than specifying a full likelihood function for the observed and missing data, we have proposed a weighted pseudo likelihood method. Using this method, a model can be built based on available data by accounting for the unobserved data via weights treated as nuisance parameters in the model. The performance of the proposed method is compared with a number of standard methods via simulation, and the method is illustrated using an example from the GenIMS study.

Respondent-Driven Sampling in an HIV at-Risk Population: The Analogy with Markov Chain Monte Carlo

✱Robert E. Weiss, University of California, Los Angeles, Department of Biostatistics, School of Public Health, Los Angeles, CA 90095-1772, *robweiss@ucla.edu*; Jason Wang, University of California, Los Angeles; Pamina Gorbach, BERG; Steve Shoptaw, University of California, Los Angeles

Key Words: Snowball Sampling, HIV, MCMC

Respondent driven sampling (RDS) is a Markov chain sampling scheme that can be compared to Markov chain Monte Carlo. Starting values, transition kernels, communication and convergence are discussed for RDS. RDS was used to sample HIV at-risk populations in Los Angeles in the SATHCAP study. We discuss some of the apparent lapses of RDS in this population in Los Angeles. In particular, the RDS sampling scheme may not be stationary, and for this particular survey study participants were heavily very low income, unemployed and homeless. An explanation is given why HIV can transmit in Markov fashion through a given population while RDS may not. We argue that prior to the start of an RDS trial it is necessary to prove that the stationary distribution of the sampling scheme is the intended population.

267 Methods for Survival Time Analyses ●

Section on Statistics in Epidemiology, Section on Nonparametric Statistics, Section on Quality and Productivity, Biometrics Section

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Semiparametric Methods for Estimating the Effect on Survival of an Experimental Time-Dependent Treatment

✱Douglas E. Schaubel, The University of Michigan, 1420 Washington Heights, Department of Biostatistics, Ann Arbor, 48109, *deschau@umich.edu*

Key Words: counterfactual, matching, non-proportionality, observational studies, propensity score, stratification

In epidemiologic studies of survival data, the treatment of interest is often time-dependent. We develop semiparametric procedures to estimate the effect of a time-dependent treatment in the presence of the following complicating factors: both an experimental and established treatment are available; pre- and post-treatment hazards are nonproportional; eligibility to receive treatment is transient; treatment assignment is not randomized. The proposed methods involve reorganizing the observed data into strata which, along with propensity score matching, set up the contrasts of interest. Treatment-by-stratum-specific survival functions are estimated nonparametrically, then appropriately averaged. Inference procedures are developed, tested via simulation, then applied to national renal failure data.

Structural Models for the Effect of a Time-Varying Exposure on Survival

✱Jessica G. Young, Harvard School of Public Health, 677 Huntington Avenue, Room 819, Boston, MA 02115, *jyoung@hsph.harvard.edu*; Miguel A. Hernán, Harvard School of Public Health; Sally Picciotto, Harvard School of Public Health; James Robins, Harvard School of Public Health

Key Words: causal effects, marginal structural models, structural nested models, g-estimation, survival analysis, inverse-probability weighting

Standard methods for estimating the effect of a time-varying exposure on survival may be biased in the presence of time-dependent confounders affected by exposure. Previously proposed methods solve this problem by estimating the parameters of marginal structural Cox models (Cox MSM) via inverse probability weighting and of structural nested accelerated failure time models (SNAFTM) via g-estimation. Here we introduce a new class of models, the structural nested cumulative failure time models (SNCFTM). We further formalize equivalencies between the SNCFTM, Cox MSM and SNAFTM. Advantages of SNCFTM over SNAFTM include smoother estimating equations in model parameters and easier accommodation of censoring; advantages of SNCFTM over Cox MSM include allowance for effect modification by time-dependent covariates. A simulation study comparing estimators from the three approaches is presented.

Joint Model of Longitudinal Binary Data with Discrete Survival Time

✱Kirsten Lum, National Institutes of Health, 6100 Executive Blvd., Room 7B03Q, Rockville, MD 20852, *lumkirst@mail.nih.gov*; Rajeshwari Sundaram, National Institutes of Health; Germaine Buck Louis, National Institutes of Health

Key Words: Survival, Likelihood, Longitudinal, Epidemiology

Time to pregnancy (TTP), the number of menstrual cycles needed to achieve pregnancy, is widely used as a measure of fecundability in epidemiologic studies. Until recently, these studies did not incorporate information on frequency and timing of intercourse with respect to the fertile days of the cycle. Dunson (2003, 2007) has shown that intercourse behavior can have a large impact on probability of conception. It is also known that intercourse behavior in couples varies greatly, as it is influenced by libido and other factors, ex. age (Wilcox, 2004). We propose a joint model for intercourse behavior using association models for binary correlated data and a discrete survival model for TTP, incorporating unobserved libido via a latent variable. We use an EM algorithm approach for the estimation procedure. These methods are applied to the NY State Angler Cohort Prospective Pregnancy study.

Semiparametric Methods for the Analysis of Clustered Survival Data from Case-Cohort Studies

✱Hui Zhang, The University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48105, *huizh@umich.edu*; Douglas E. Schaubel, The University of Michigan; John Kalbfleisch, The University of Michigan

Key Words: case-cohort study, clustered data, Cox model, estimating equation, robust variance, survival analysis

Case-cohort sampling is often employed in large cohort studies. Although most existing methods apply to univariate survival data, subjects are often clustered in epidemiologic studies. We propose estimating equation based methods for three case-cohort designs that are applicable to clustered survival data. A marginal model is assumed, with a baseline hazard and regression coefficient that are common across clusters. The proposed regression parameter and cumulative hazard estimators are consistent and asymptotically normal, with variances that can be estimated empirically. The proposed procedures have increased efficiency relative to some existing methods. The proposed

estimators are evaluated analytically, investigated through simulation, and then applied to a study of mortality among Canadian dialysis patients.

A Method for Finding the Nadir of Nonmonotonic Relationships

✱Fei Tan, Florida State University, 314 Pennell Circle, Apt 6, Tallahassee, FL 32310-5546, *ftan@stat.fsu.edu*

Key Words: J-shaped relationship, nadir estimation, Cox model, polynomial splines, free-knot

Different methods have been proposed to model the J-shaped or U-shaped relationship between a risk factor and mortality so that the optimal risk-factor value (nadir) associated with the lowest mortality can be estimated. The basic model used is the Cox Proportional Hazards model. Here we consider polynomial splines with free knots, a natural generalization of the quadratic, the change point and the fixed-knot spline methods. They can be applied to risk factors that do not have a good Normal transformation. Asymptotic normality and consistency of the maximum partial likelihood estimators are established under certain condition. The new method is motivated by and applied to the nadir estimation in nonmonotinic relationships between BMI (body mass index) and all-cause mortality. Its performance is compared with that of existing methods.

Exponential Tilt Models in the Presence of Censoring

✱Chi Wang, Johns Hopkins University, 615 N. Wolfe St., Baltimore, MD 21205, *chwang@jhsph.edu*; Zhiqiang Tan, Rutgers, The State University of New Jersey; Thomas A. Louis, Johns Hopkins University

Key Words: Exponential tilt model, Censored data, Nonparametric likelihood

Various semi-parametric models can be used to describe survival distributions in randomized clinical trials. Assuming a constant change in hazards for the placebo and treatment groups leads to proportional hazard models. Alternatively, assuming a scale change in survival functions leads to accelerated failure time models. In this paper, we consider exponential tilt models (ETM) that assume a parametric form for the density ratio of the two survival distributions, and develop a nonparametric likelihood method for estimation in the presence of censoring. ETM includes a broad array of parametric models such as lognormal model and gamma model. Although exponential tilt models in the absence of censoring have been widely studied in connection with case-control studies and biased sampling our work presents a first step in extend applications of such models for survival analysis.

A Multivariate Classification Tree for Survival Data with Competing Risks

✱Fiona M. Callaghan, University of Pittsburgh, 1931 W Russet Ct, Apt 10, Appleton, WI 54914, *fmc2@pitt.edu*

Key Words: competing risks, classification and regression tree, survival tree, non parametric, multivariate, cumulative incidence function

Classification trees are the most popular tool for categorizing individuals into groups and subgroups based on particular outcomes of interest. To date, trees have not been developed for the competing risk situation where two or more competing risk outcomes are of equal interest to the researcher. In this paper we propose a classification tree to analyze survival data with multiple competing risk outcomes using a composite measure of discrimination between subgroups. The method used for growing and pruning the tree is described and its performance is demonstrated with simulations in a variety of competing risk model configurations. The multivariate competing risk tree is compared to currently available tree-based methods. Finally, the new method is illustrated with an example from the liver transplant waitlist registry.

268 Solving Statistical Problems in Marketing Science ●

Section on Statistics and Marketing

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Data Mining Pitfalls: The Nontechnical Killer Errors

✱Sam Koslowsky, Harte Hanks, 55 Fifth Avenue, New York, NY 10003, *sam_koslowsky@harte-hanks.com*

Key Words: database marketing, predictive modeling, data mining, marketing objectives

Most marketers and analysts agree that data mining plays a critical role in assuring a successful campaign. At times, errors, in executing an analytics project can subtly creep in. While analytic and technical errors can certainly harm a data mining exercise, most of the problems that emerge in a predictive modeling project have little to do with technical issues. Rather, basic reasoning, and marketing related issues are at fault. While technical issues can also cause havoc in an analysis, most difficulties and problems emanate from the mundane issues. What is the correct marketing objective? Is the right analysis being applied? Are my files being prepared diligently? Is the sample valid? Are the tools appropriate? Does it make sense? Perhaps these standard issues are not as routine as we think?

Combined Personalization Approach to Different Stages of Marketing

✱Hye-Kyoung Kim, Intellidyn Corp., 105 Maxess Road, Suite 121, Melville, NY 11747, *hkim@intellidyn.com*; Dmitri V. Kuznetsov, Intellidyn Corp.

Key Words: marketing, personalization, lead incubation, web analytics, customer interactions

By types of interactions with prospects and customers, the direct marketing technologies can be splinted in four groups/stages: (1) Initiation of interactions (by mass-media/direct/tele/e-marketing), (2) Real-time web interactions, (3) Long-term interactions (or Lead incubation) and (4) After event/transaction interactions. Each of these groups can be considered separately or like a joined up to 4-courses scheme. We consider a methodology to apply personalization approaches across all these groups/stages to create a full service cycle. Performance of each stage and heredity between stages can highly benefit not only from accumulated analysis of interaction activities with prospects (while very important) but also very significantly from comprehensive and up-to-date transaction, credit, demographic and behavioral databases, providing "atomic" or clustering level of data analytics.

Opinion Spreading and Personalization in Marketing

✱Dmitri V. Kuznetsov, Intellidyn Corp., 175 Derby Street, Unit 40, Hingham, MA 02043, *dkuznetsov@intellidyn.com*

Key Words: sociophysics, mediaphysics, word of mouth, opinion spreading, marketing, personalization

Opinion spreading (by word of mouth including internet opinion exchange and propagation) is not only one of the key factors for marketing, it is the factor which transforms effects of other (advertising and economic) factors making them less or more significant, or even critical for abrupt transitions to a new market share or product/service demand. Such effects cannot be estimated using typical statistical techniques like, for example, a regression dependency between factors and target value. Using our statistical mediaphysics methodology we were able to uncover many challenging marketing issues and propose ways how they can be put under control to a better marketing performance. To describe a big picture we combined mass media/direct/tele/e-marketing with personalization approach. These aspects of marketing represent a part of the corresponding methodologies at Intellidyn Corp.

Customer Segmentation: Data Mining Application for the Auto Insurance Industry

✱David S. Dobson, North Carolina State University, 2586 East 17th Avenue, Vancouver, BC Canada, *david.dobson1@gmail.com*

Key Words: Automobile Insurance, Cluster Analysis, Customer Data, Customer Profiling, Customer Segmentation, Data Mining

As the auto insurance industry becomes more value-competitive, insurance companies are facing challenges to maintain their market share. To understand their customers better and create effective marketing initiatives, companies can profile their customers using data mining techniques. Using customer profiling, companies can initiate customized marketing messages that reach out to the target group by tailoring products and services to the individual customer's needs. This can lead to increased customer engagement and stronger customer relationships. This paper demonstrates the steps that are taken in the data mining process. Fictitious data are used as input data for this study. The research goal is to segment current customers based on their driver history data using cluster analysis, and then to profile each segment to understand which are high risk customers and which are safe.

Statistical Assessment of the Sample Size Requirement for Projection of the Automobile Sales Volume in the U.S. Market

✱Keiko I. Powers, J. D. Power and Associates, 324 Hickory Grove Drive, Thousand Oaks, CA 91320, *keiko.powers@jdpa.com*

Key Words: sample size requirements, confidence intervals, negative binomial distribution, sales volume projection, US automobile market

The paper illustrates how a confidence interval approach based on the negative binomial distribution is used for determining the sample size requirement for projecting the automobile sales in the U.S. market. For the sales projection, it is important to know the relationship between the number of automobile dealers in the sample and the accuracy of the estimated sales volume. The sales volume trend is very volatile, and therefore it is clear that a relatively large number of dealers' data are needed for accurate sales projection. However, because the overall sales volume varies considerably among dealers, statistically deriving sample size requirements called for a new methodology. The paper provides a step-by-step description of the approach and demonstrates its performance by using actual vehicle point-of-sales data collected from several thousand dealers in the U.S. market.

A System of Models for the Process of Automobile Sales Behavior

✱James G. Wendelberger, Urban Science Applications, Inc., 620 Los Pueblos, Los Alamos, NM 87544, *jgwendelberger@urbanscience.com*

Key Words: spatial, marketing, advertising, automotive, inventory, profitability

The main driver in automobile sales behavior is location. A microlocation gravity model is used to represent consumer buying behavior as a function of spatial location. The model includes sub-models with the ability to adjust theoretical advertising spend, outlet locations and theoretical inventory or to change the area geography or attraction and draw to determine the model profitability and sales. The model uses agent based modeling to simulate the automobile purchase process. This model uses "StarLogo" software from the MIT. Sub-models include: buyer location distribution, random area geography, the "S" curve of the effect of advertising spend, the "S" curve as the effect of inventory, the distance metric between buyers and the sales outlets, a spatial interaction type attraction and draw model, a multinomial distribution of error model as well as various profitability models.

Local Regression with Random Censored Data for Customer Wallet Data Mining

S. Tom Au, AT&T Labs - Research; ✱William Pepe, AT&T Labs - Research, *wp8082@att.com*

Key Words: customer wallet, censored data, local regression, data mining, share-of-wallet, prediction

Most corporations record each customer's spending for a product and are interested in estimating each customer's related total spending (customer wallet). One approach is to identify key predictors (business size, industry, etc) that drive the spending and develop a quantile regression to estimate the wallet. The quantile is chosen to normalize the corporation's overall market share. This approach is unable to provide measure of uncertainty and wallet may be less than spending. Here, we develop a framework using local regression with random censored data. Using local regression to model the total spending and the predictors, assume the corporation's share of wallet follows a Beta distribution. Under certain conditions, we can estimate the parameters of the Beta distribution, the local regression, and the conditional distribution of the total spending of each customer.

269 Technology for Teaching Statistics in the Traditional and Online Classrooms ●

Section on Statistical Education

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Tablet PC Applications in Statistics Education, Part I

✱Christopher R. Bilder, University of Nebraska-Lincoln, Department of Statistics, 340 Hardin Hall North, Lincoln, NE 68583, *chris@chrisbilder.com*; Christopher J. Malone, Winona State University

Key Words: class capture, computer, OneNote, podcast, PowerPoint, Word

This is the first of two presentations on the use of Tablet PCs in statistics education. In this presentation, the main focus will be on learning environments where the instructor primarily is the only one with a Tablet PC. The presentation will demonstrate basic applications of Tablet PCs, such as their use in lectures and for grading. More advanced applications will also be demonstrated including their use in capturing classroom content when paired with recording software.

Tablet PC Applications in Statistics Education, Part II

✱Christopher J. Malone, Winona State University, P.O. Box 5838, Winona, MN 55987, *cmalone@winona.edu*; Christopher R. Bilder, University of Nebraska-Lincoln

Key Words: Tablet PCs, Classroom Presenter, Class capture, Clicker, Dyknow, Computer

This is the second of two presentations on the use of Tablet PCs in statistics education. In this presentation, the main focus will be on learning environments where both students and the instructor have Tablet PCs. Student-to-student and instructor-to-student interaction is enhanced though the use

of Tablet PCs. Tablet PC applications allow students to effectively complete homework assignments and allow for computer screen output to be shared and annotated interactively during class. In addition, Tablet PCs can serve as an alternative to using "clickers" in the classroom that are used for obtaining immediate feedback from students.

Look, Ma, No Textbook! Computerized Statistics Learning

✱Sandra Clarkson, Hunter College of CUNY, 695 Park Avenue, New York, 10065, *sclarkso@hunter.cuny.edu*; Bill Williams, Hunter College of CUNY

Key Words: computer-based, understanding, multimedia, statistics

An innovative multimedia program ActivStats allows students access to film clips, interactive text and a statistical data analysis package. It allows instructors to teach in a totally different manner—without a text. Our lectureless course serves 800 students a semester in more than 30 sections taught in computer labs by adjunct instructors. Students study statistics using the CD-ROM, discuss the material with their professors and with each other, and use a creative, intuitive data analysis program to graph data and calculate statistics. Graphics is emphasized. As part of the course requirement, students are given moderately large data sets and asked to analyze them and write a 'report' to a non-statistician. For assessment, we are using the CAOS pretest and posttest to assess growth in understanding and a common final. We will discuss our philosophy, approach and results.

Expect More, Get More: The Joys of Teaching Online

✱Olcay Akman, Illinois State University, Department of Mathematics, Normal, IL 61790-4520, *oakman@ilstu.edu*

Key Words: self-learning, mastery-based

An introductory-level statistics course designed for nonmajors in the general education program was taught online based on a self-learning structure and a mastery-level assessment set up. In this talk we will describe the contents, the unique structure and the instructor's non-traditional role in this course. A comparison of student performance with those of the traditionally taught section along with some actual student comments will be shared.

Assessment of Online Statistics Courses

✱David W. Stockburger, U.S. Air Force Academy, 215 Bent Nail Way, Monument, CO 80132, *DStockburger@MissouriState.edu*; S. David Kriska, The Ohio State University

Key Words: on-line, education, statistics, assessment

Two forces, the emphasis on deep learning of knowledge, skills, and abilities and the ability of students to select courses from a variety of sources, will define online statistics courses in the future. This poses dual questions of how can colleges and universities assess the quality of these courses relative to their stated goals and outcomes and how can students best select courses to meet their needs. This paper presents recommendations for the assessment of online statistics courses that focus on the acquisition of statistical thinking and course quality in contrast to coverage of content and student satisfaction.

The Use of Microsoft Excel for Statistical Analyses: An Update

✱John D. McKenzie, Jr., Babson College, Math/Science Division, Babson College, 231 Forest Street, Babson Park, MA 02457-0310, *mckenzie@babson.edu*

Key Words: algorithms, graphical displays, missing data, software documentation, spreadsheet, statistical software

At the 2001 Joint Statistical Meeting Jon Cryer and the author organized a well-attended panel on the use of Microsoft Excel for statistical analyses. The panelists explained that while Excel, the most used spreadsheet package in the world, was being used to perform statistical analyses at schools and in the workplace, there were some major concerns about using Excel for instruction and research. In the last seven years Microsoft has released Excel 2003 and Excel 2007. This paper will present what has changed and what has remained the same with the statistical capabilities of Excel since that panel presentation in New York City.

270 Topics in Physical and Engineering Sciences ●

Section on Physical and Engineering Sciences

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

The Importance of Identifying Different Components of Mixture Distribution in Reliability Predictions

✱Yili Hong, Iowa State University, 305 Snedecor Hall, Ames, IA 50011, *hong@iastate.edu*; William Q. Meeker, Iowa State University

Key Words: Hazard rate, maximum likelihood, regression analysis, transformer maintenance, transformer reliability, Weibull

It is well known that a mixture of two exponential distributions with different hazard rates will always have a decreasing hazard rate. Relatedly, data from a mixture of distributions with two increasing hazard rates can behave like from a distribution with decreasing hazard (e.g., reliability predictions based on data from a mixture of units with two or more different physical designs could be seriously wrong if the pooled data are used). Thus, it is important to identify components of the mixture and do statistical inference based on the stratified data. In this paper, the importance of this principle is illustrated through the power transmission transformer lifetime data. Comparisons are made between the predictions based on the pooled data and stratified data and the importance of correct stratification in practice is shown. Some suggestions for practitioners are also given.

Generalized Inferences for the Common Scale Parameter of the Several Pareto Populations

✱Sumith Gunasekera, University of Nevada, Las Vegas, Department of Mathematical Sciences, University of Nevada, Las Vegas, NV 89154, *sg@unlv.nevada.edu*; Malwane Ananda, University of Nevada, Las Vegas

Key Words: Generalized p-values, Generalized tests, Pareto distribution, Common scale parameter

The Pareto distribution is widely used to model data relevant to social, scientific, geophysical, and many other types of observable phenomena. A problem of interest in this article is statistical inferences concerning the common scale parameter of several Pareto distributions. Using the generalized p-value approach, exact confidence intervals and the exact tests for testing the common scale parameter are given. Examples are given in order to illustrate our procedures. A simulation study is given to demonstrate the performance of the proposed procedures.

On Classification of Bivariate Distributions Based on Mutual Information

✱Mohamed Habibullah, Northeastern University, 2115 Huntington University, Boston, MA 02115, *m.habibullah@neu.edu*; Mohammad Ahsanullah, Rider University

Key Words: Mutual information, entropy, survival distribution, bivariate distributions

Among all measures of independence between random variables, mutual information is the only based on information theory (Cover and Thomas) . Mutual information takes into account all kinds of dependencies between variables (i.e., both the linear and nonlinear dependencies). In this paper, we have classified some well-known bivariate distributions into two classes of distributions based on their mutual information. The distributions within each class have the same mutual information. These distributions have been used extensively as survival distributions of two component systems in reliability theory.

Inference on Multiresolution Spatial-Temporal Process with Application

✱Wanli Min, IBM T. J. Watson Research Center, 1101 Kitchawan Road, #26-204, Yorktown Heights, NY 10598, *wanlimin@us.ibm.com*

Key Words: Saptial-Temporal, Mean-Field, Kriging

We consider a spatial stochastic process where the spatial measurements are taken at the two extreme settings: measurements are taken with high spatial resolution only once per year; measurements are taken at very limited fixed spatial locations continuous over time. The goal is to predict the dependent observable at any spatial location and time. We establish an inference procedure based on mean-field approximation and kriging with drift taking into account the effect of exogenous environmental variables on the observables. We also propose an algorithm of sequential design by adding new sampling points to existing layout. Application to a real dataset is reported in the end.

Analyzing Hybrid Randomized Response Data with a Binomial Selection Procedure

✱Feiyi Jia, Oakland University, Department of Mathematics and Statistics, Rochester, MI 48309, *fjia@oakland.edu*; Gary C. McDonald, Oakland University

Key Words: Contaminated data, Indifference-zone selection procedure, Probability of a correct selection, Sample size determination, Operating characteristics

The operating characteristics (OCs) of an indifference-zone ranking and selection procedure are derived for randomized response binomial data. The OCs include tables and figures to facilitate tradeoffs between sample size and a stated probability of a correct selection, i.e., correctly identifying the binomial population (out of) characterized by the largest probability of success. Measures of efficiency are provided to assist the analyst in selection of an appropriate randomized response design for the collection of the data. A hybrid randomized response model, which includes the Warner model and the Greenberg et al. model, is introduced to facilitate comparisons among a wider range of statistical designs than previously available. An example comparing failure rates of contraceptive methods is used to illustrate the use of these new results.

On Diffusion Processes with Gaussian Marginals

✱Rabab S. Jafri, University of Nottingham, 29-Boundary Road,Beeston,, Nottingham, NG9 2RG United Kingdom, *pmxrs@nottingham.ac.uk*; Sergey Utev, University Of Nottingham,UK

Key Words: Gaussian marginals, Characterizing equation, Foster-Lyapunov criteria, Dynkin's formula

I'm going to talk about the characterization of the homogeneous one-dimensional diffusion processes given that the marginals are Gaussian. Then I'll explain how to construct a peculiar diffusion process which does not satisfy usual linear growth conditions.

271 Advances in Functional and Related Data Analysis ▲

Section on Nonparametric Statistics, Section on Statistical Computing, IMS

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Functional Singular Value Decomposition and Applications

✱Wenjing Yang, University of California, Davis, One Shields Ave, Dept. of Statistics, Davis, CA 95616, *wjyang@ucdavis.edu*; Hans G. Müller, University of California, Davis

Key Words: Functional data analysis, Functional correlation, Sparse data, Covariance estimation

Inspired by the canonical expansion of compact operators in functional analysis, we develop the concept of functional singular value decomposition (fSVD). It is shown that both functional singular values and singular component functions can be consistently estimated even when data are observed sparsely and with measurement error as is the case in many longitudinal studies. Applications of fSVD include functional partial least squares and functional correlation. Due to the need to invert compact operators, some correlation measures in FDA such as functional canonical correlation are plagued by numerical instability. We apply fSVD to derive a stable functional correlation measure which is shown to be stable, and demonstrate its performance in simulations and applications.

Manifold Models for Functional Data

✱Dong Chen, University of California, Davis, Department of Statistics, 1 Shields Avenue, Davis, CA 95616, *dchen@wald.ucdavis.edu*; Hans G. Müller, University of California, Davis

Key Words: functional manifold, warping, Isomap, structural mean

In the situation where functional data scatter around a nonlinear manifold in function space, the usual sample mean and linear functional principal component representation are not effective. We demonstrate this for time-warped functional data, and develop a modified unwrapping method, based on a version of ISOMAP (Tenenbaum 2000). Our modification is geared towards the case of noisy data by changing the distance on which the core multidimensional scaling step is based. We demonstrate in simulations and applications that the nonlinear representations resulting from unwrapping functional data indeed are of interest.

BLUP for Functional Data

✱Ana Kupresanin, Arizona State University, *maria@mathpost.asu.edu*

Key Words: BLUP, Canonical correlation, Functional data, Reproducing kernel Hilbert space, Stochastic process

We consider the regression setting where either the predictor or the response or both are random functions defined on a compact subset of R. We approach the regression problem from the best linear predictor point of view and our motivation comes from the fact that regression is related to canonical correlation analysis. The resulting form for the best linear unbiased predictor is derived using the isomorphism that relates a second-order process to the reproducing kernel Hilbert space generated by its covariance kernel. We implement the proposed predictor and study it through simulation.

A Generalized CP Criterion for Evaluating Derivative Estimates

✱Richard Charnigo, University of Kentucky, Department of Statistics, Patterson Tower 851, Lexington, KY 40506-0027, *RJCharn2@aol.com*; Cidambi Srinivasan, University of Kentucky

Key Words: nonparametric regression, local modeling, global modeling, smoothing

A challenging problem in functional data analysis is to estimate a mean response and its derivatives simultaneously. Charnigo and Srinivasan recently proposed a new "compound estimator" for this purpose. However, the compound estimator (or any other nonparametric estimator that might be employed) requires the user to specify tuning parameters. Selection of the tuning parameters is difficult in this context because familiar criteria like CP and BIC reflect the quality with which the mean response has been estimated; if estimation of the derivatives is equally (or more) important, one should explicitly take into account not only how well the mean response has been estimated but also how well the derivatives have been estimated. In this talk we describe a new "generalized CP criterion" that meets this need, and we examine its performance on simulated and real data.

Improved Covariance Inversion and Derivative Estimation for Sparse and Irregularly Observed Functional Data

✱Bitao Liu, University of California, Davis, Department of Statistics, Davis, CA 95616, *btliu@wald.ucdavis.edu*; Hans G. Müller, University of California, Davis

Key Words: Functional Data Analysis, Longitudinal Data, Inverse Covariance, Derivatives

We discuss an improved functional principal component analysis approach to represent sparsely observed longitudinal trajectories. It is based on advances in estimating the error variance and shrinkage which in turn leads to better estimates of inverses of covariance matrices of the underlying stochastic process. The gains of improved covariance inversion are demonstrated for the difficult problem of estimating derivatives from sparse and irregularly observed data. We present some theoretical results for this problem, simulation comparisons and data illustrations.

A Distribution-Free Test of Constant Mean in Linear Mixed Effects Models

✱Xinlei Wang, Southern Methodist University, Dallas, TX 75275, *swang@smu.edu*; Johan Lim, Yonsei University

Key Words: mixed effects model, permutation test

We propose a distribution free procedure, an analogy of the DIP test in nonparametric regression, to test whether means of responses are constant over time in repeated measures data. Unlike existing tests, the proposed procedure requires very minimal assumptions to the distributions of both random effects and errors. We study the asymptotic reference distribution of the test statistic analytically and propose a permutation procedure to approximate the finite-sample reference distribution. The size and power of the proposed test are illustrated and compared with competitors through several simulation studies. We find that it performs well for data of small sizes, regardless of model specification. Finally, we apply our test to a data example to compare the effect of fatigue in two methods used for cardiopulmonary resuscitation.

Robust Curve Classification

Hyejin Shin, Auburn University; ✱Nedret Billor, Auburn University, 221 Parker Hall, Auburn, AL 36849, *billone@auburn.edu*

Key Words: functional data, classification, outliers

Modern data collection methods due to the rapid technological advances are returning observations that should be viewed as a result of digitized recording from continuous curves. It is well known that the direct use of traditional multivariate techniques for classification to such data often fails and several new statistical methodologies have been proposed for the last few decades to consider the nature of such data. However, many of those techniques assume that the dataset is free of outliers, which is never true in practice, thus are highly sensitive to outliers. In this talk, we present the effects of outliers in classification of curves and propose a new classification technique based on robust functional PCA and statistical data depth to minimize the effects of outliers. The practical usefulness of the proposed method will be demonstrated by simulation and a real-data application.

272 Nonparametric and Semiparametric Bayesian Methods ●

Section on Bayesian Statistical Science, Section on Nonparametric Statistics, Section on Statistical Computing

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Bayesian Clustering with Regression

✱Fernando A. Quintana, Pontifícia Universidad Católica de Chile, Departamento de Estadística, Av. V. Mackenna 4860, Macul, Santiago, International RM22 Chile, *quintana@mat.puc.cl*; Peter Mueller, M.D. Anderson Cancer Center; Gary L. Rosner, University Texas MD Anderson Cancer Center

Key Words: Product Partition Models, Dirichlet Process, Nonparametric Bayes, Survival Curve, Posterior Predictive Inference

We propose a model for covariate-dependent clustering (i.e., we develop a probability model for random partitions that is indexed by covariates). The motivating application is inference for a clinical trial. As part of the desired inference we wish to define clusters of patients. Defining a prior probability model for cluster memberships should include a regression on patient baseline covariates. We build on product partition models (PPM). We define an extension of the PPM to include the desired regression. This is achieved by including in the cohesion function a new factor that increases the probability of experimental units with similar covariates to be included in the same cluster. We discuss implementations suitable for continuous, categorical, count and ordinal covariates.

A Mixture of Semiparametric Models

✱Yeonok Lee, The University of Iowa, Iowa City, IA 52242, *yelee@stat.uiowa.edu*

Key Words: semiparametric model, mixture model

A model is developed that allows full flexibility on location and scale, producing estimates of entire conditional distribution as a mixture of semiparametric models. The state probabilities are permitted to depend on both parametric and nonparametric components. All parameters and a unknown univariate smooth function are accessed in a state space model framework, requiring the number of operations proportional to the sample size in Markov Chain Monte Carlo algorithm. Two illustrations are presented. In the first example, quantiles of distribution of an Engel curve is studied. In the second illustration, we investigate the inequality of earnings of men conditional on age, education attainment (high school graduate and university graduates), and year (1977 and 1994) by means of quantiles of the posterior conditional distributions.

Nonparametric Bayesian Models with Various Base Measures

✱Ju Hee Lee, The Ohio State University, 1958 Neil Avenue Cockins Hall, Room 404, Columbus, OH 43210, *juheelee@stat.osu.edu*

Key Words: Multivariate normal distribution, nonparametric bayesian models, shrinkage estimators

I address the problem of estimating the mean of a multivariate normal distribution as follows: From unknown distribution G there arise independently and identically distributed but unobserved draws, theta_1,..., theta_n in a set \Theta. Given these \theta's, conditionally independent variables y_1,..., y_n are observed, where the distribution of y_i given \theta_i is N(\theta_i, \sigma^2). One reasonable model for this problem is a nonparametric Bayesian model with various base measures. Candidates for the base measure include a normal distribution, a t distribution or an improper distribution. The nonparametric Bayesian model probabilistically partitions the data coordinates and shrinks them into a common center within a partition. We can compare these nonparametric Bayesian models to each other as well as to other shrinkage estimators such as the James-Stein estimator.

Bayesian Semiparametric Modeling and Functional Data Analysis with Application to the Texas Lottery

✱Anandamayee Majumdar, Arizona State University, *ananda@math.asu.edu*; Randall Eubank, Arizona State University

Key Words: curve registration, functional data analysis, posterior inference, predictive distribution

State lotteries employ sales projections to determine appropriate advertised jackpot levels for some of their games. This paper focuses on prediction of sales for the Lotto Texas game of the Texas Lottery. A novel prediction method is developed in this setting that utilizes functional data analysis concepts in conjunction with a Bayesian paradigm to produce predictions and associated precision assessments.

Bayesian Joint Modeling Method for Survival Analysis with Curve Predictors

Key Words: Markov Chain Monte Carlo, Proportional Hazards

We propose a relatively simple semi-parametric joint model using spline basis, in which the usage of the splines simplifies the parameterizations and the joint modeling framework allows synergistic benefit between the regression of functional predictors and proportional hazards modeling of survival data. Additional advantage is that our model handles survival analysis that is rather complex involving irregularly and sparsely sampled curves with correlation within each curve.

The Probit Stick-Breaking Process Mixture for Conditional Distribution Modeling with Variable Selection

✱Yeonseung Chung, The University of North Carolina at Chapel Hill, NC 27517, *chungy@email.unc.edu*; David Dunson, National Institute of Environmental Health Science

Key Words: Conditional distribution estimation, Kernel stick-breaking processes, Mixture of experts, Hypothesis testing, Stochastic search variable selection

This article considers methodology for flexibly characterizing the relationship between a response and multiple predictors. Goals are (1) to estimate the conditional response distribution with the distributional changes across the predictor space, and (2) to identify important predictors for the response distribution change both with local regions and globally. We propose a probit stick-breaking process mixture (PSBPM) of normal regressions for the conditional response distribution. A stochastic variable selection (SSVS) structure is incorporated in the PSBPM for both regression coefficients and predictor-dependent mixture weights, which facilitates a global variable selection. Relying on the conditional distribution estimates at different predictor points, we perform local variable selection. The methods are illustrated through simulation and applied to an epidemiologic study.

Using Prior Information in Bayesian Nonparametric Models

✱David B. Dahl, Texas A&M University, Department of Statistics, Texas A&M University, College Station, TX 77843-314, *dahl@stat.tamu.edu*

Key Words: Clustering algorithm, Dirichlet process mixture model, Dirichlet process prior, Gene annotation, Gene expression, Nonuniform pairwise priors

Integration of data from several sources and technologies is a burgeoning field in bioinformatics. Some data naturally lead to formal statistical models, yet others may merely convey proximity among observations. In the context of clustering, methods are often either model-based or distanced-based. In many cases, however, both types of information are available. I propose a hybrid approach that is simultaneously model-based and distance-based. Specifically, I show how the usual Dirichlet Process mixture model framework can be adapted to incorporate pairwise distances between observations. One application area is incorporating gene annotation information in statistical models for gene expression. Another application is protein structure prediction, where one can estimate protein torsion angles distributions using both (phi,psi) angle pairs and RMSD distances from peptides.

273 Government Statistics

Business and Economics Statistics Section, Social Statistics Section

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Frame Improvements to Statistics Canada Business Register

✱Yanick Beaucage, Statistics Canada, 11-A RHCoats Bldg, Tunney's Pasture, Ottawa, ON K1A 0T6 Canada, *yanick.beaucage@statcan.ca*; Wesley Yung, Statistics Canada

Key Words: Business Register, Register Redesign, Centralization of data

The Statistics Canada Business Register has been recently redesigned and as part of the redesign a number of initiatives regarding frame improvement have been developed. This presentation will discuss improvements that were made in several areas. These include the birthing and deathing processes, the increased availability of administrative data, the industrial coding of units and improvements made to the automatic coding environment. Also the presentation will discuss the centralization of data, the process being put in place to adapt to survey needs and the new approach to update the frame. Finally, it will discuss how some data are being verified for consistency and how month to month changes to BR data are reported and monitored.

The Quality Assurance Survey and Its Impact on the Canadian Business Register

✱Tyler Kirkland, Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *tyler.kirkland@statcan.ca*; Zeeshan Mahmood, Statistics Canada; Caroline Pelletier, Statistics Canada

Key Words: quality assurance survey, business register, statistics canada, NAICS, industrial coding, establishments

The Business Register (BR) is used by most business surveys at Statistics Canada as their sampling frame. The key variable on the BR is the industrial code that Statistics Canada assigns for all Canadian businesses based on their main business activity descriptions. Proper industrial coding is crucial to ensure surveys minimize costs and sustain high-quality respondent relations. Developed as part of the BR redesign, the Quality Assurance Survey (QAS) is an on-going monthly survey designed to measure the proportion of incorrectly coded establishments as well as the proportion of dead establishments. This paper will present the methodology and implementation for the QAS as well as some preliminary results. In addition, it will discuss the important role the survey plays at Statistics Canada along with presenting the current and future uses of the survey.

An Analysis of Key Differences in Micro Data: Results from the Business List Comparison Project

✱Lucia Foster, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *Lucia.S.Foster@census.gov*; Kristin Fairman, Bureau of Labor Statistics; C. J. Krizan, U.S. Census Bureau; Ian Rucker, Bureau of Labor Statistics

Key Words: business register, business list, micro data

The Bureau of Labor Statistics and the Bureau of the Census each support a business register, a universe of all U.S. business establishments, created from independent sources. The lists serve a number of critical functions including supplying aggregate data inputs for certain national accounts generated by the Bureau of Economic Analysis. Elvery et al (2006) found a high level of consistency between the two lists at the micro-level but also identified key differences. We build upon the prior research by utilizing a more recent reference year to address certain key micro-level differences highlighted by Fixler and Landefeld (2006). We address wage and employment differentials in Washington State, in Texas's headquarters and oil and gas industries, and in California's computer and electronic product manufacturing industry. We also provide an overview of differences between the lists.

Is the Grass Always Greener on the Other Side? Assessing the Determinants of Individual Well-Being Across Europe

Luisa Corrado, University of Cambridge; ✱Aqib Aslam, University of Cambridge, Faculty of Economics and Politics, Sidgwick Avenue, Cambridge, CB39DD United Kingdom, *maa49@econ.cam.ac.uk*

Key Words: Multilevel Modelling, Regional Well-Being, Social-Exchange, Social-Distance, Inter-Personal Factors

We use multilevel modeling to judge whether location and neighborhood effects play a significant role in determining the well-being of individuals across the EU. The key concept is that there are certain features of social position (or 'location') that affect subjective well-being. As we have hierarchical data grouped by neighborhoods, we have assessed the determinants of well-being using a multilevel modeling approach. Empirical findings support the idea that well-being is a multidimensional phenomena affected both by objective indicators such as income, socio-demographic factors, geographic and institutional characteristics of the country where each individual lives, and by more subjective measures such as personal and relational factors, as well as by favorable comparisons with reference groups.

Multiple Imputation Approaches for Right-Censored Wages in the German IAB Employment Register

✱Thomas Buettner, Institute for Employment Research, Regensburger Str. 104, Nuremberg, International 90480 Germany, *thomas.buettner@iab.de*; Susanne Rässler, Otto-Friedrich University Bamberg

Key Words: multiple imputation, missing data, censored wage data, simulation study

Right-censored wages are very common with administrative data from social security systems like the IAB employment register, which is based on the data of the German unemployment insurance. In order to be able to analyze wages with this register, the German Institute for Employment Research (IAB) is using multiple imputation approaches to impute the missing wage information by draws of a random variable from a truncated distribution based on Markov chain Monte Carlo techniques. In addition, we suggest a new multiple imputation method based on GLS estimation which does not presume homoscedasticity of the residuals. In a simulation study, we use uncensored wage information from an income survey (German Structure of Earnings Survey, GSES) to compare different imputation approaches in order to confirm the necessity as well as the validity of the new approach.

Disclosure Protection: A New Approach to Cell Suppression

✱Bei Wang, U.S. Census Bureau, 4600 Silver Hill rd., Suitland, MD 20746, *bei.wang@census.gov*

Key Words: Cell Suppression, Linear Programming, Network Flow

Census products and related programs use cell suppression to protect data that is sensitive to our respondents. A disclosure procedure is applied before any data goes out for publication. The underlining algorithm used is a network flow model. We will review the disclosure procedure and how well the model does. A question that always arises is "how is a near optimized solution to be determined to the Cell Suppression Problem (CSP)?" A new linear programming approach is used in this research. The algorithm is applied to Survey of Business Owners (SBO)'s Hispanic data and comparisons with the 2002 publications are made.

Remeasuring the Size of the World's Economy

✱Frederic A. Vogel, The World Bank, 10109 Waterside Drive, Burke, VA 22015, *fvogel@msn.com*

Key Words: Purchasing Power Parities, International Comparison Program, Gross Domestic Product

The final results of the International Comparison Program show a 40 percent reduction in the sizes of many of the world's largest economies such as China and India compared to previous measures. The purpose of the ICP is to provide currency converters in lieu of exchange rates to convert national statistics such as the Gross Domestic Product into a common currency. Purchasing Power Parities were computed for 146 countries following a complex process that began with each country providing prices for a well-defined set of more than 1,000 goods and services. The resulting estimates of Purchasing Power Parities provide the capability to compare economic aggregates across countries that are independent of exchange rates. The ICP became one of the world's largest statistical initiatives involving 146 countries and represented a re-engineering of much of the statistical methodology.

274 Procedures for Nonparametric Regression Modeling

Biometrics Section, Section on Nonparametric Statistics

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Covariate Adjusted Correlation Analysis with Application to FMR1 Premutation

✱Danh V. Nguyen, University of California, Davis, 1 Shields Ave, MS1C, Davis, CA 95616, *ucdnguyen@ucdavis.edu*; Damla Senturk, The Pennsylvania State University

Key Words: Binning, Conditional correlation, Fragile X syndrome, Local method of moments, Nonparametric partial correlation, Pearson correlation

Motivated by molecular data on female premutation carriers of the fragile X mental retardation 1 (FMR1) gene, we present a new method of covariate adjusted correlation analysis to examine the association of messenger RNA (mRNA) and number of CGG repeat expansion in the FMR1 gene. The proposed adjusted correlation uses local conditional correlations, which are local method of moments estimators, to estimate the Pearson correlation between two variables, adjusted for a third observable covariate (activation ratio, AR). The final covariate adjusted correlation estimator is shown to be consistent. Application to our data on 165 female carriers indicates that the association between mRNA and CGG repeat after adjusting for AR is stronger and is more in-line with association seen in male carriers. The results support a joint additive and multiplicative effect of AR in the literature.

Self-Modeling Regression with Regression Splines and Random Curve-Specific Parameters: Application to Arterial Pulse Pressure Waveforms

✱Lyndia C. Brumback, University of Washington, Box 357232, Seattle, WA 98115, *LynB@u.washington.edu*; Douglas Tommet, Institute for Aging Research; Richard Kronmal, University of Washington

Key Words: functional data, nonlinear mixed effects models, pulse waveforms, self-modeling

Self modeling regression is a method for analyzing sets of observed curves. It is based on the relatively simple assumption that the x and y axes can be separately transformed in a parametric manner for each curve so that the data from all curves lie approximately on one typical curve. When the typical curve is modeled with a regression spline and the curve-specific transformational parameters are modeled as random (Normal with mean zero), the model may be under-parameterized and the variance components may be estimated poorly. Simulations show that a random effects distribution that forces the realized curve-specific transformational parameters to have mean zero or the inclusion of a fixed transformational parameter improves estimation. The methods are applied to arterial pulse pressure waveform data from the Multi-Ethnic Study of Atherosclerosis.

Using Bernstein Polynomials To Model Misclassification in BI-RADS Breast Density Measurements

✱Charlotte C. Gard, University of Washington, Department of Biostatistics, Box 357232, Seattle, WA 98195, *gardc@u.washington.edu*; Elizabeth R. Brown, University of Washington

Key Words: breast density, misclassification, Bernstein polynomials

Breast cancer risk prediction models that include breast density (BD) measured on the four-point American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) scale have been shown to have slightly better predictive accuracy than models that do not include BD. Because BI-RADS BD is subject to misclassification, it is thought that inclusion of BD measured continuously could further improve the predictive accuracy of the models. Unlike BI-RADS BD, however, continuous BD is not routinely measured in clinical practice. With a final goal of predicting continuous BD for women for whom it is not available, we take a Bayesian approach to modeling the misclassification in BI-RADS BD. Continuous BD is modeled using Bernstein polynomial priors. We allow for radiologist-specific misclassification in BI-RADS BD and linkage of this misclassification to radiologist characteristics.

A Varying-Coefficient Model for the Evaluation of Time-Varying Concomitant Intervention Effects in Longitudinal Studies

✱Colin O. Wu, National Heart, Lung, and Blood Institute, 2 Rockledge Center, 6701 Rockledge Drive, Room 9212, Bethesda, MD 20892-7913, *wuc@nhlbi.nih.gov*; Xin Tian, National Heart, Lung, and Blood Institute; Heejung Bang, Weill Medical College of Cornell University

Key Words: Change-Point Models, Concomitant Intervention, Longitudinal Study, Polynomial Splines, Shared Parameter Model, Varying-Coefficient Model

Concomitant interventions are often introduced during a longitudinal clinical trial to patients who respond undesirably to the prespecified treatments. Because the initiation of a concomitant intervention may depend on the patient's general trend of pre-intervention outcomes, regression approaches that treat the presence of the intervention as a time-dependent covariate may lead to biased estimates for the intervention effects. We propose a varying-coefficient mixed-effects model to evaluate the patient's longitudinal outcome trends before and after the patient's starting time of the intervention. Our model leads to less biased estimates of the intervention effects. Nonparametric estimation and inferences of the coefficient curves and intervention effects are developed using B-splines. Our methods are demonstrated through a longitudinal clinical trial in depression and heart disease.

Inference on Quantile Regression for Heteroscedastic Mixed Models

✱Huixia Wang, North Carolina State University, 2501 Founders Drive, 209F Patterson Hall, Raleigh, NC 27695, *wang@stat.ncsu.edu*

Key Words: Apnea duration, Estimating equation, Longitudinal data, Rank score test, Weighted quantile regression

This paper develops two weighted quantile rank score tests for the significance of fixed effects in a class of mixed models with nonhomogeneous groups. One test is constructed by simply weighting the residuals to account for heteroscedasticity, while the other test is based on asymptotically optimal weights accounting for both heteroscedasticity and correlation. Without appropriate weights to account for heteroscedasticity, the quantile rank score tests often perform poorly. In finite samples, the test with optimal weights tends to provide marginal improvements over the one with simpler weights unless the intra-subject correlation is extremely high. We illustrate the value of the proposed methods by modeling several quantiles of the apnea duration of elderly during normal swallowing.

Varying Coefficient Model for Epigenetic Modifications

✱Dong Wang, University of Nebraska-Lincoln, 340 Hardin Hall North, UNL East Campus, Lincoln, NE 68583, *dwang3@unl.edu*

Key Words: ChIP on chip, tiling microarrays, varying coefficient model, epigenetics, histone modification

ChIP-chip is a powerful tool for epigenetic research. However, current statistical methods are developed primarily for detecting transcription factor binding sites, and there is currently no satisfactory method to incorporate covariates such as time, hormone levels, and genotypes. In this study, we develop a varying coefficient model for epigenetic modifications such as DNA methylation and histone acetylation. By taking into account the special features of ChIP-chip data, a plug-in type method is derived for bandwidth selection in the local linear fitting of the varying coefficient model. Our results show that analyses using the proposed varying coefficient model can effectively detect diverse characteristics of epigenetic modifications over genomic regions as well as across different treatment conditions.

Nugget Estimation for a Class of Nonparametric Semivariograms

✱Jeffrey Spence, The University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-8896, *Jeffrey.Spence@UTSouthwestern.edu*; Patrick Carmack, The University of Texas Southwestern Medical Center; Qihua Lin, The University of Texas Southwestern Medical Center; Richard Gunst, Southern Methodist University; William R. Schucany, Southern Methodist University

Key Words: brain imaging, B spline, fMRI, kriging, spatial modeling, SPECT

Kriging estimators and their variances depend importantly on the behavior of the semivariogram near the origin. It is often the case that the semivariogram is discontinuous there when a nugget effect is present. Shapiro and Botha (1991) have offered a flexible class of nonparametric estimators based on a spectral representation of the covariance, but it necessarily extrapolates to the origin. In this paper we have incorporated an additive shift parameter in the nonlinear optimization routine used by Cherry et al. (1996) and estimated this feature by minimizing ISE with a spline-smoothed empirical estimate of the semivariogram. In simulation these estimates have significantly lower MSE compared to the parametric estimates from the same realizations. An application to functional brain imaging data is featured to demonstrate the automation advantages for multiple regional spatial modeling.

275 Case-Control Studies and Logistic Regression: New Extensions

Biometrics Section, Biopharmaceutical Section, Section on Statistics in Epidemiology

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Optimum Two-Stage Sampling with Inexpensive Error-Prone and Expensive Error-Free Measures for Designing a Case-Control Study

✱Jun-mo Nam, National Cancer Institute, Executive Plaza South/ Romm 8028, Biostatistics Branch, DCEG, 6120 Executive Blvd, Rockville, MD 20852-7244, *namj@mail.nih.gov*

Key Words: cost efficient double sampling, fallible classification, gold standard, optimum sample size, validation sub-sampling

We consider a case-control study design using two-stage sampling in which case and controls are classified as exposed or not-exposed by an inexpensive but error-prone measure in the 1st stage and it's sub-sampling is classified by the expensive but error-free measure in the 2nd stage. We provide the optimum proportion for validation size to the primary one which maximize precision of estimation of a true odds ratio for a given total cost, and present the optimum sample size formulae for designing a case-control study. We compare our optimal design with McNamee's optimal designs. Examples are given for numerical illustrations. A gain in precision of estimation and cost reduction resulted from the optimal allocation in a comparison with those not using the optimal rule is substantial when a unit cost ratio of the gold standard to a fallible classification is very high.

Bias Reduction and a Solution for Separation of Logistic Regression with Missing Covariates

✱Tapabrata Maiti, Iowa State University, 50014, *taps@iastate.edu*; Vivek Pradhan, Cytel Inc.

Key Words: Small-sample, Bias, EM-algorithm, likelihood, separation

Logistic regression is an important statistical procedure used in many disciplines. The standard software packages for data analysis are generally equipped with this procedure; where the maximum likelihood estimates of the regression coefficients are obtained iteratively. It is well known that the estimates from the analyses of small or medium sized samples are biased. Also in finding such estimates, often a separation is encountered in which the likelihood converges but at least one of the parameter estimates diverges to infinity. Standard approaches of finding such estimates do not take care of these problems. Moreover, the missingness in the covariates adds an extra layer of complexity to the whole process. In this article we address these three practical issues—bias, separation, and missing covariates—by means of simple adjustments.

Assess the Effect of Measurement Error in Estimating a Logistic Regression Model

✱Huey-Miin Hsueh, National Cheng-Chi University, Department of Statistics, 64, Chih-Nan Road, Sec. 2, Wenshan, Taipei, International 11623 Taiwan, *hsueh@nccu.edu.tw*; Chih-Yi Chien, National Cheng-Chi University; Yuan-chin I. Chang, Institute of Statistical Science, Academia Sinica

Key Words: Logistic regression, Measurement error

This work aims to assess the effects of measurement error of a covariate variable in estimation of a logistic regression model. Begg and Lagakos (1992) considered a uniform measurement error and showed that the MLE converges to true value almost surely and has asymptotic normality. However, a nonzero bias in the asymptotic distribution is induced by the measurement error. Here we find a sufficient condition such that the two MLEs, either based on mismeasured or accurate data, are asymptotically equivalent in terms of consistency and asymptotic normality. The measurement error have negligible effect in estimation. The required measurement error rate is nonincreasing as the sampling order. We further propose a design to generate observations with required magnitude of measurement error. Numerical studies are conducted to justify the theoretical results.

Statistical Inferences for Outcome-Dependent Sampling Design with Multivariate Outcomes

✱Tsui-Shan (Eva) Lu, The University of North Carolina at Chapel Hill, CB#7420, School of Public Health,, 3105 McGavran-Greenberg Hall, Chapel Hill, NC 27599, *tlu@bios.unc.edu*; Haibo Zhou, The University of North Carolina at Chapel Hill

Key Words: Outcome-dependent sampling, Multivariate, Empirical, Semiparametric

An outcome-dependent sampling (ODS) (Zhou et al. 2002) scheme is a useful sampling scheme where one observes the exposure with a probability, maybe unknown, depending on the outcome. However, the ODS design and statistical inference procedures with multivariate situations still remain undeveloped. We discuss a general sampling design and inference methods using ODS under multivariate settings. The proposed estimators are consistent and more efficient than those obtained using a simple random sample of the same size. The estimators are semiparametric in nature that all the underlying distributions of covariates are modeled nonparametrically using the empirical likelihood methods. The multivariate-ODS design provides a cost-effective approach to further improve study efficiency.

Association Models for Clustered Data with Binary and Continuous Responses

✱Lanjia Lin, Florida State University, *lanjia@stat.fsu.edu*; Stuart R. Lipsitz, Harvard Medical School; Dipankar Bandyopadhyay, Medical University of South Carolina; Debajyoti Sinha, Florida State University

Key Words: bivariate discrete and continuous responses, bridge distribution, probability integral transformation, clustered data

We consider the study in which the sampling unit is a cluster of subjects and each member of the cluster has mixed bivariate responses, a binary and a continuous outcome. We propose a new bivariate random effects model which induces associations between the binary outcomes in the cluster, the continuous outcomes in the cluster, the binary and continuous outcomes from different subjects in the cluster, as well as the binary and continuous outcomes from the same subject in the cluster. For the ease of interpretation of the regression effects, the marginal model of the binary response probability, integrated over the random effect, preserves the logistic form, and the marginal regression function of the continuous response preserves the linear form. Finally, we illustrate the maximum likelihood estimation of parameters of our model via the analysis of data from a development toxicity study.

Fitting Stratified Proportional Odds Models by Amalgamating Conditional Likelihoods

Bhramar Mukherjee, The University of Michigan; ✱Jaeil Ahn, The University of Michigan, M4153, 1420 Washington Heights, Ann Arbor, 48109-2029, *jaeil@umich.edu*

Key Words: conditional likelihood, nuisance, proportional odds, sandwich estimate

Classical methods for fitting a varying intercept logistic regression model to stratified data are based on the conditional likelihood principle to eliminate the stratum-specific nuisance parameters. However, classical conditioning techniques do not apply to the general K-category cumulative logit model (K>2) with varying stratum-specific intercepts and there is no reduction due to sufficiency; the nuisance parameters remain even in the conditional likelihood. We propose a methodology to fit stratified proportional odds model in a general regression set-up, by amalgamating conditional likelihoods. We provide a robust sandwich estimate of the variance of the proposed estimator. Simulation studies comparing the proposed method with a random effects model on the stratification parameter are also furnished.

Exact Trend Tests for Correlated Categorical Data

✱Chris Corcoran, Utah State University, 3900 Old Main Hill, Logan, UT 84322-3900, *corcoran@math.usu.edu*; Pralay Senchuadhuri, Cytel Inc.

Key Words: categorical data, clustered data, exact test

We extend a previously developed exact trend test for correlated binomials in order to compare clustered ordinal multinomials. We also suggest a modification that allows within-cluster variation with respect to the investigative factor (e.g., dose group), when analyzing either correlated binomials or multinomials. We use examples to illustrate the method and its implementation in the StatXact software package.

276 Multiple Comparisons in Clinical Trials ●▲

Biopharmaceutical Section, Biometrics Section

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

On the Generalization of Closed Testing Procedures with Null Hypotheses Grouped into Families

✱Hanjoo Kim, University of Pennslyvania School of Medicine, Department of Biostatistics, 423 Guardian Drive, Blockley Hall 501, Philadelphia, PA 19104, *hanjoo@mail.med.upenn.edu*; A. Richard Entsuah, Merck Research Laboratories; Justine Shults, University of Pennsylvania

Key Words: Closure method, Multiple testing procedure, Familywise error rate

We present a method of devising sequentially rejective multiple testing procedures with strong control of the familywise error rate for testing null hypotheses that are grouped into families. The method is a generalization of the closure method which provides a unified framework for constructing closed testing procedures in numerous applications. In particular, the method is applied to gatekeeping problems, in which hypotheses in a family can be tested only if hypotheses in the previous family have been rejected, to derive a class of stepwise gatekeeping procedures that are highly useful in clinical trials. We illustrate the computation of the adjusted p-values using a real clinical trial example which compares venlafaxine versus lithium monotherapies in patients with Bipolar II major depressive episode.

Resampling-Based Measures for Understanding the Nature of Treatment Effects of Multiple Endpoints in Clinical Trials

✱Mushfiqur M. Rashid, U.S. Food and Drug Administration, Division of Biometrics IV/OB/OTS/CDER, 10903 New Hampshire Ave, WO 22-6238, Silver Spring, MD 20993, MD 20993, *rashidm@cder.fda.gov*; Mohammad F. Huque, U.S. Food and Drug Administration

Key Words: Re-sampling Method, Marginal Model, Conditional Model, Multiple Endpoints, Association Measures, PTE measure

In a clinical trial treatment effects for endpoints can be either of overlapping or nonoverlapping nature, in the sense that certain single or a group of multiple endpoints is able to explain or not able to explain a part or whole of the treatment effect of another endpoint. This information can be useful in assessing the total benefit of a treatment for a set of multiple endpoints of a trial. An easy to understand measure for this purpose is the proportion of treatment effect (PTE) of a clinical endpoint explained by the treatment effects of other endpoint(s) of interest. Conventionally, it is estimated by the ratio of two statistics. However, this ratio estimate has been statistically challenging. It can produce a wide confidence interval beyond [0, 1] interval.

Type I Error Rate Adjustment for Superiority Test Conditional on Establishment of Noninferiority in Clinical Trials

✱Jiacheng Yuan, Medtronic CardioVascular, 3576 Unocal Place, Santa Rosa, CA 95403, *jiacheng.yuan@medtronic.com*; Tiejun Tong, University of Colorado, Boulder

Key Words: non-inferiority, superiority, type I error rate adjustment, conditional test, clinical trial

In clinical trials, it is often desirable to test superiority conditional on establishment of noninferiority. Committee for Proprietary Medicinal Products issued guidelines accepting that no type I error rate adjustment is required for this switch. However, a researcher pointed out that switching between noninferiority and superiority without any adjustment reduces to simultaneous testing of both, which not only has a logic flaw similar to post-hoc specification of the null hypothesis but also leads to a higher chance for erroneous claiming superiority compared with only one null hypothesis to be tested. We suggest an adjustment which makes the second step test depend on the first step test, hence it is a conditional test in the real sense and the switch does not reduce to simultaneous testing. Also our suggested adjustment decreases the chance of erroneous claiming superiority.

Power and Type I Error Comparisons for Seven Multiplicity Adjustment Methods

✱Jin Xu, Merck & Co., Inc., 351 N. Sumneytown Pike, PO Box 1000, North Wales, PA 19454, *jin_xu@merck.com*; Li David, Merck & Co., Inc.; Kenneth Liu, Merck & Co., Inc; Ivan S.F. Chan, Merck Research Laboratories; Jie Chen, Merck & Co., Inc.

Key Words: Multiplicity adjustment, Power, Type I error

Many new multiplicity adjustment procedures have been proposed in the past two decades including methods proposed by Holm, Hochberg, Hommel, Rom, Dunnet and Tamhane, and step-up bootstrap and permutation methods. The last two methods do not require any distribution assumptions. Most statisticians realize the necessity to adjust for multiple hypothesis tests in a clinical trial. The question is which adjustment method ensures control of the Type I error and maximizes the power of the clinical trial. We will expand the research by Dunnet and Tahmane (1993) by evaluating the power of the seven widely used multiplicity adjustment methods listed above under various data structures, including negatively correlated data. We will also complement their research by evaluating the Type I error of these methods for both the one-sided or two-sided testing scenarios.

Estimation After Step-Down Tests in the Dose-Finding Studies

✱Yi-Hsuan Tu, National Cheng Kung University, Department of Statistics, No. 1 Ta-Hsueh Road, Tainan, International 70101 Taiwan, *yhtu@stat.ncku.edu.tw*; Ying Kuen Cheung, Columbia University

Key Words: Dose finding, Familywise error, Multiple comparisons

This presentation discusses the estimation of treatment effect at the identified minimum effective dose (MED) in a dose finding placebo-controlled study. The identification of MED is often formulated as a multiple comparison problem, and as such, demands a strict control of the familywise error rate (FWE). While the Bonferroni procedure and the Dunnett's procedure control the FWE, step-down tests may generally improve the probability of selecting the correct dose by avoiding multiplicity adjustment. In order to preserve the nominal coverage probability, confidence lower bound for the treatment effect at the identified MED is equal to the hypothesized treatment effect. We propose a two-stage procedure in which the step-down test is conducted at a slightly more conservative level than nominal, so that the remaining error rate may be spent in the estimation stage.

Reverse Fixed-Sequence Procedures in Clinical Trials

✱Alex Dmitrienko, Eli Lilly and Company, 10913 West 144th Street, Overland Park, KS 66221, *alexei@lilly.com*; Suktae Choi, Eli Lilly and Company; Ajit C. Tamhane, Northwestern University; Don Johns, Eli Lilly and Company

Key Words: Multiple comparisons, Clinical trials, Familywise error rate

This talk discusses multiple comparison procedures based on multiple sequences of nested hypotheses (reverse fixed-sequence procedures). Within each sequence, hypotheses are tested sequentially and each one is tested if the previous one is accepted (if a hypothesis is rejected, the remaining hypotheses are automatically rejected). Problems of this kind are encountered in clinical trials with two-dimensional objectives that include multiple dose-control tests and multiple testing strategies (superiority and noninferiority tests). Reverse fixed-sequence procedures control the familywise error rate with respect to multiple hypotheses in each sequence and multiple sequences of hypotheses. A two-stage reverse fixed-sequence procedure for a clinical trial with dose-control comparisons including superiority and noninferiority tests will be used to illustrate this methodology.

Partitioning K Treatments According to Multivariate Equivalence with Respect to a Standard: The Common Covariance Case

✱Weixing Cai, Syracuse University, Department of Matheamtics, 215 Carnegie Hall, Syracuse, NY 13244, *wcai01@syr.edu*; Pinyuen Chen, Syracuse University

Key Words: Multivariate bioequivalence, Multivariate normal, Mahalanobis distance, Optimal procedure, Correct decision, Sample size determination

Multivariate bioequivalence has been considered in biopharmaceutical fields. We use the multivariate normal model and assume the covariance matrices are equal. We consider the Mahalanobis distance as a metric to measure the dissimilarity between a treatment and a standard and define the equivalence region to be a fixed ellipsoid. Our goal is to classify k treatments into two disjoit subsets according to their multivariate equivalece to the standard. Partitioning k treatments is reduced to partitioning k noncentral chi-square or noncentral F distributions. Optimal partitioning procedures are proposed with controlling the probability of making a correct decision. The problem of sample size determination is addressed. We used simulation examples and a real data example to show that our selection procedures work well.

277 Inference Using Simulation

Section on Statistical Computing, Section on Nonparametric Statistics
Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Saddlepoint-Based Bootstrap Inference for Quadratic Estimating Equations

✱Robert Paige, Texas Tech University, 79409, *r.paige@ttu.edu*

Key Words: bootstrap, saddlepoint approximations

We propose an easy to implement method for making small sample inference about the root of an estimating equation expressible as a quadratic form in normal random variables. It is based on saddlepoint approximations to the distribution of the estimating equation whose unique root is a parameter's maximum likelihood estimator (MLE), while substituting conditional MLEs for the remaining (nuisance) parameters. A key result of Daniels (1983) enables us to relate these approximations to those for the estimator of interest. The proposed method is equivalent to a parametric bootstrap percentile approach where Monte Carlo simulation is replaced by saddlepoint approximation. It finds applications in many areas of statistics including, nonlinear regression, time series analysis and inference on ratios of regression parameters in linear models.

Applying Resampling To Analyze the Sensitivity of a Hypothesis Test to Confounding

✱William M. Goodman, University of Ontario Institute of Technology, 2000 Simcoe Street North, Oshawa, ON L1H 7K4 Canada, *bill.goodman@uoit.ca*

Key Words: resampling, simulation, confounders, sensitivity analysis

This paper demonstrates a resampling-based method for sensitivity analysis, to quantify the risk that a hypothesis test result may be in error, as a consequence of a confounder's influence. Due to confounding, test statistics apparently on the tail of the null distribution, and so "significant," may in fact be in non-critical regions-if the null distribution is corrected to reflect the confounder's impact on sample measurements. The proposed method is analogous to creating a power curve, based on the varying risks of Type II error depending on the relative proximity of the true population mean to the null mean. Similarly, we display varying risks of reaching false positive conclusions, due to confounding, depending on the relative severity of the bias. The method is presented in the context of a specific case; then the paper explores how the technique can be generalized.

On Markov Chains with Periodic Transition Probabilities

✱Rajeeva L. Karandikar, Cranes Software Internatonal Limited, Shankaranarayan Building, 4th floor, 25 M G Road, Bangalore, 560001 India, *rkarandikar@gmail.com*

Key Words: Markov chains, periodic, classification, hidden Markov model, simulation, genes

Non-stationary Markov chains have received very little attention in the literature. We consider chains where the transition probabilities from time n to (n+1) are periodic in n with period p. Such chains occur in genetics—a gene can be modeled as a k-step Markov chain with transition probabilities that have 3-periodicity. We consider the problem: given a realization of a (finite state space) stochastic process to classify it as arising from a Markov chain with periodic transition probabilities with a period p or from a Markov chain with stationary transition probabilities. We propose a statistic and illustrate its efficacy via simulation results. We will also consider the same question for a hidden Markov model (HMM) when the driving chain has periodic transition probabilities.

Bootstrap Bandwidth Selection for Estimating a Conditional Probability

✱Sundar Subramanian, New Jersey Institute of Technology, Department of Mathematical Sciences, University Heights, Newark, NJ 07102, *sundars@njit.edu*

Key Words: Inverse probability of nonmissingness weighted estimator, Least squares cross validation, Missing at random, Pilot bandwidth, Survival analysis

Data-drivenbandwidthsforcomputingsub-densityfunctionkernelestimators are proposed. The bandwidths are obtained as minimizers of estimates of the mean integrated squared error (MISE). The smoothed bootstrap offers a motivation for choosing the MISE estimate for minimization. The efficacy of the proposed procedure is investigated through simulation studies. Illustrations are provided in the context of survival function estimation from right-censored data with missing censoring indicators.

Using Bootstrap To Construct Adjusted Likelihood-Based Confidence Intervals

✱Esa I. Uusipaikka, University of Turku, Department of Statistics, Assistentinkatu 7, Turku, International 20014 Finland, *esa.uusipaikka@utu.fi*; Kaisu Ikonen, University of Turku; Anna Wiksten, University of Turku

Key Words: Bootsrap, Confidence interval, Profile likelihood

The most often used method to construct confidence intervals is delta method the special case of which is Wald method. Profile likelihood-based confidence intervals generally have, however, coverage probabilities that are better in agreement with nominal level. Also sometimes delta method gives intervals containing impossible values, but this never happens with profile likelihood-based intervals. In 1996 a new method to calculate approximate profile-likelihood based confidence intervals was introduced. This method is applicable in any parametric statistical model to any smooth enough parameter function, not only to components of the parameter vector. In this paper bootstrap method is used to adjust limits of profile likelihood-based confidence interval in a new way that is well suited with the method of calculation.

Relative Entropy Measures of Asymmetry with Applications

✱Bhaskar Bhattacharya, Southern Illinois University Carbondale, Carbondale, IL 62901, *bhaskar@math.siu.edu*

Key Words: categorical, mirror image, skewness, central asymmetry, Kullback-Leibler divergence, coefficients

We propose a measure of asymmetry of a probability density function (PDF) by considering the relative entropy between itself and its (appropriately defined) mirror image. The measure is shown to be useful for detecting asymmetry in distributions of categorical or continuous random variables. Asymmetries of a PDF near its center and away from the center are also investigated. Applications are discussed for two-way tables and in linear regression models. Monte Carlo simulations show that the proposed measures/tests have good size and power properties when compared with competitors, even for smaller samples. Two illustrative examples are analyzed.

Simulation Studies for Mixture Transition Distribution Model in High-Order Markov Chains

Din Chen, South Dakota State University; ✱Y. L. Lio, University of South Dakota, 414 East Clark Street, Department of Mathematical Scieces, Vermillion, SD 57069, *Yuhlong.Lio@usd.edu*

Key Words: Markov Chains, High-order temporal dependence, Mixture transition distribution, Maximum likelihood estimation

The mixture transition distribution (MTD) model was developed by Raftery to model high order Markov chains. This MTD model significantly reduced the number of parameters to be estimated in modeling high order Markov chains, but it was still flexible to represent the temporal dependence of the high order Markov Chains. Difficulties arouse in the estimation of the model parameters with the nonlinear constraints from MTD specification in addition to local optima. A transformation is proposed to make the nonlinear constraint into box-constraints for global optimization. Simulation studies are conducted to compare and demonstrate the implementations of the transformation. Finally, the proposed transformation is also applied to fit the MTD model to a real data (song of weed Pewee).

278 Invited Poster Presentations: Clinical Trial Designs

Biometrics Section

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

Sequential Continual Reassessment Method for Two-Dimensional Dose Finding

✱Ying Yuan, M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Unit 447, Dept. of Biostatistics, Houston, TX 77030, *yyuan@mdanderson.org*; Guosheng Yin, The University of Texas M.D. Anderson Cancer Center

Key Words: clinical trial, drug combination, schedule

It is common to encounter two-dimensional dose finding in phase I trials, for example, in trials combining multiple drugs, or in single-agent trials that simultaneously search for the maximum tolerated dose and treatment schedule. We propose a simple and adaptive two-dimensional dose-finding design that can accommodate any type of single-agent dose-finding method. In particular, we convert the two-dimensional dose-finding trial to a series of one-dimensional dose-finding subtrials along shortened line search segments by fixing the dose level of one drug, and conduct them sequentially. Based on the MTD obtained from the completed one-dimensional trial, we eliminate the doses that lie outside of the search range based on the partial order, and thereby adaptively and efficiently reduce the two-dimensional dose-finding space. We illustrate the design through extensive simulation studies.

Optimal Adaptive Randomized Designs for Clinical Trials

✱Yi Cheng, Indiana University South Bend, 1700 Mishawaka Ave., South Bend, IN 46634, *ycheng@iusb.edu*

Key Words: Balanced design, Clinical Trials, Decision theory, Multiple arms, Utility, Randomized sequential allocation

Optimal decision-analytic designs are deterministic. Such designs are appropriately criticized in the context of clinical trials because they are subject to assignment bias. On the other hand, balanced randomized designs may assign an excessive number of patients to a treatment arm that is performing relatively poorly. We propose a compromise between these two extremes, one that achieves some of the good characteristics of both. We introduce a constrained optimal adaptive design for a fully sequential randomized clinical trial with k arms and n patients. An r-design is one for which, at each allocation, each arm has probability at least r of being chosen. An optimal design among all r-designs is called r-optimal. We also show that, in a trial with an r-optimal design, fewer patients are assigned to an inferior arm than when following a balanced design.

Three-Dose Cohort Designs in Cancer Phase I Trials

✱Bo Huang, Pfizer, Inc., 175 Hawthorne Drive Apt 25J, New London, CT 06320, *bo.huang@pfizer.com*

Key Words: cancer phase I trials, CRM, maximum tolerated dose, multiple dosing

Traditional designs for phase I clinical trials assign the same dose to patients in the same cohort. In this paper, we present a new class of designs for cancer phase I trials which initially rapidly escalate by allowing multiple doses (usually 3) to be assigned to each cohort of patients. The class of designs, called the LMH-CRM by administering different percentiles of the maximum tolerated dose (MTD), denoted 'low,' 'medium,' 'high,' is proven to be consistent and coherent (a commonsense property of phase I trials for dose escalation and de-escalation). Three designs (slow, moderate and fast) are derived based on different dose-escalation restrictions. Simulation results show that moderate and fast LMH-CRM combine the advantages of the CRM with one patient per cohort and three patients per cohort: it accurately estimates the MTD; controls overall toxicity rates; and is time efficient.

Dose-Finding in Phase I Clinical Trials Based on Toxicity Probability Intervals

✱Yuan Ji, M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Unit 237, Houston, TX 77030, *yuanji@mdanderson.org*

Key Words: Dose escalation, Dose de-escalation, MTD, Trial monitoring, Bayesian design

Most phase I clinical trials conducted at the M. D. Anderson Cancer Center use the algorithmic 3+3 design, despite the availability of more advanced model-based designs such as the continual reassessment method. Through simple statistical modeling and computing, we develop a dose-finding design that can be easily understood and implemented by nonstatisticians. We propose a beta/binomial Bayesian model and a probabilistic up-and-down rule that allow all possible dose-assignment actions to be tabulated in a spreadsheet. We have developed an Excel macro (available at http://odin.mdacc.tmc.edu/~yuanj/software.htm) that generates trial monitoring tables, which contain the dose-assignment actions corresponding to various toxicity outcomes. The new design outperforms the 3+3 design and performs comparably to other model-based methods in the literature.

Identifying Optimal Cumulative Treatment Regimes in Early-Phase Clinical Trials

✱Thomas M. Braun, The University of Michigan, 1420 Washington Heights, M4063 SPH II, Department of Biostatistics, Ann Arbor, MI 48109, *tombraun@umich.edu*

Key Words: dose-finding, Phase I trial, maximum tolerated dose, Bayesian methods, adaptive design

Optimal treatment regimes for cancer are no longer based upon a single administration of an experimental agent, as the cumulative dose needed to maximize cytotoxicity must be divided into a series of administrations to prevent treatment-related toxicity to patients. However, the optimal combination of (1) number of administrations, (2) timing of each administration, and (3) dose given at each administration is typically unknown. We present a dynamic outcome-adaptive design that simultaneously optimizes both dosing and administration in order to determine a maximum tolerated treatment regime. We describe our modeling approaches and graphically present the execution of our design in a hypothetical trial in 60 bone marrow transplant patients designed to select which of 12 possible regimes leads to toxicity in approximately 30% of patients within 116 days of beginning treatment.

279 Contributed Poster Presentations

Biometrics Section, Biopharmaceutical Section, Business and Economics Statistics Section, ENAR, IMS, Section on Bayesian Statistical Science, Section on Health Policy Statistics, Section on Statistics and Marketing, Section on Statistics in Epidemiology, Section on Survey Research Methods

Tuesday, August 5, 10:30 a.m.–12:20 p.m.

A Longitudinal Study of Nigerian Stock Prices

✱Hamadu Dallah, University of Lagos, Faculty of Acturial Science and Insurance, Faculty of Business Administration, Lagos, 9600001 Nigeria, *adeleke22000@yahoo.com*; Ismaila A. Adeleke, University of Lagos

Key Words: Nigerian Stock market, Time series Analysis, ARIMA models, ARCH, GARCH, Lognormal and Gamma distributions

Stock options are ways in which investors can manage the various risk level of their portfolios. Usually, this class of derivatives consists of agreements to buy or sell shares of stock at a prescribed time in the future; hence, determining their fair market value requires prediction of future prices of shares. Identifying probability distribution that characterizes the stock prices has been extensively investigated in empirical finance and related fields. This paper models the behavior of stock prices traded on the floor of Nigerian Stock Exchange using time series analysis and parametric distribution models. Contrary to expectation of conditional volatility models requiring ARCH or GARCH methods, most Nigerian stock prices exhibited a simple Autoregressive process (AR(1)). Moreover, it is quite interesting to notice that these stock prices can be modeled using lognormal and gamma.

The Transformed GARCH Model for the Asymmetric Volatility

In-Kwon Yeo, Sookmyung Women's University; ✱Juyeon Park, Sookmyung Women's University, Chungpa-dong 52, Yongsan-gu, Seoul, International 140-742 South Korea, *larc7221@hanmail.net*

Key Words: asymmetric volatility, GARCH model, Yeo-Johnson transformation, Bayes factors, MCMC

One of main characteristics of financial time series is that the volatility is rarely constant and usually asymmetric. The asymmetry in volatility means that responses to bad or good news are different, especially more sensitive to unfavorable factors. ARCH and GARCH model are well known models for modeling the conditional heteroscedastic effect. However, these models don't allow for asymmetric effects. Modified models such as GJR and TGARCH have been developed to handle them. In this paper, we introduce a new type of GARCH model in which the error terms in the conditional variance are transformed via Yeo-Johnson transformation and then are squared. The proposed model can be employed to detect asymmetric news effect. Parameters are estimated within Bayesian framework and Bayes factors are computed to compare the performance of existing models and the proposed model for real data.

Statistics for Comparing Two Treatments with Placebo, with Selection of Better Treatment

✱Xianhuang Zhou, Bristol-Myers Squibb Company, 777 Scudders Mill Road, Plainsboro, NJ 08536, *xianhuang.zhou@bms.com*

Key Words: ranking and selection, two-stage design, sequential design, mean and variance, bias correction

Two-stage designs are widely used in clinical trials, usually with a selection procedure in the first stage and hypothesis test at the end of the second stage. This topic comes from the clinical trial QALS, which is a two-stage, phase II, randomized clinical trial of CoQ10 against placebo. The first stage will choose between two doses of CoQ10 to see which one is better and the second stage is to test the better dose group of CoQ10 vs. placebo using observations from both stages. We will first introduce some estimators for this design, and derive their exact mean and variance formulas which are based on normal assumption. We will also study the asymptotic behavior of these variances. Based on the means and variances of these statistics, the associated test statistics will be constructed.

Exploring the Benefits of Adaptive Sequential Designs in Time-to-Event Endpoint Settings

✱Sarah Emerson, Stanford University, Department of Statistics, 334 Olmsted Road, Apt. 333, Stanford, CA, *semerson@stanford.edu*; Scott S. Emerson, University of Washington; Kyle Rudser, University of Washington

Key Words: Clinical trials, drug discovery

We consider a family of adaptive designs that allow sample size readjustment based on the results of the first interim analysis, and we match each adaptive design to a standard group sequential design that has an equivalent power function. We investigate and compare the trade-offs between efficiency (as measured by number of observed events required) and cost (a function of the number of subjects accrued and length of observation) for a variety of patient accrual and trial cost scenarios. In particular, we consider the relative contribution of per-patient costs and the calendar time costs of financing the drug development process.

Multiple Imputation of Ordinal and Count Outcomes in a Multiple Sclerosis Clinical Trial Using Data at Dropout

Peter B. Imrey, The Cleveland Clinic and Case Western Reserve University; ✱John Barnard, The Cleveland Clinic, 9500 Euclid Avenue/Wb4, Cleveland, OH 44195, *barnarj@ccf.org*; Matthew Karafa, The Cleveland Clinic; the Avonex Combination Trial (ACT) Investigators, The Cleveland Clinic Foundation

Key Words: clinical trials, missing at random, multiple sclerosis, multiple imputation, proportional odds model, negative binomial regression

Modern missing data methods—including multiple imputation (MI), pattern mixture models, selection models when data are missing at random (MAR), and analyses of sensitivity to nonrandom dropout (MNAR)—have been used infrequently for clinical trial analysis. We illustrate MI by predictive mean matching in a 2x2 factorial trial of therapies for relapsing-remitting multiple sclerosis. Multiple imputations were generated for proportional odds analysis of the ordinal primary outcome, (categorized) new and/or enlarging T2-hyperintense MRI lesions, and negative binomial regression analysis of relapse counts. Imputation models merged available data at dropout with imputed data between dropout and target follow-up time.

A Dunnett-Bonferroni–Based Parallel Gatekeeping Procedure for Dose-Response Clinical Trials with Multiple Endpoints

✱Isaac Nuamah, Johnson & Johnson Pharmaceutical R&D, LLC, 1125 Trenton-Harbourton Road, Titusville, NJ 08560, *inuamah@prdus.jnj.com*; Haiyan Xu, Johnson & Johnson Pharmaceutical R&D, LLC; Jingyi Liu, University of California, Davis; Pilar Lim, Johnson & Johnson Pharmaceutical R&D, LLC

Key Words: Dunett-based test, parallel gatekeeping, dose-response, clinical trials, multiple endpoints, Bonferroni

The issue of multiple testing in studies with multiple endpoints and multiple active doses arises commonly in dose-response clinical trials. We propose a Dunnett-Bonferroni based parallel gatekeeping procedure in a dose-response clinical trial with multiple endpoints. It follows the Dunnett-based parallel gatekeeping strategy of Dmitrienko et al. (Pharmaceut. Statist. 2006; 5: 19-28), but differs in the calculation of critical values. The Dunnett-based parallel gatekeeping procedure relies on some assumptions that could be difficult to justify in typical clinical trials, and might be challenged by regulatory agencies. Our proposed method avoids these assumptions. While it is potentially less powerful than a Dunnett-based procedure, we show that the power loss is very minimal. In most cases this method is easier to implement compared to the Dunnett-based procedure.

Parameterization of CRM Model

✱Jianfen Shu, University of Virginia, POBOX 800717, Charlottesville, VA 22908, *jshu@virginia.edu*

Key Words: Phase I Dose-finding, MTD, CRM, Clincal Trial

The continual reassessment method (CRM) was first proposed by O'Quigley, Pepe and Fisher to challenge the standard designs in cancer studies and has gained great popularity. The CRM is designed to estimate the highest dose whose toxicity rate is closet to the target MTD. A simple one-parameter mathematical working model has been shown sufficient to appropriately locate the right dose. Some researchers use two-parameter logistic model to estimate the toxicity rates for the doses under study. However, theoretically, model-based approaches for dose-finding designs concentrate patients on a single dose. Therefore, there isn't enough information from data to estimate two parameters. We will show the one-parameter model is more stable than the two-parameter one through simulation.

Design Strategies for a Proof-of-Concept Study with Futility Analysis and Go/No-Go Criteria

✱Yong-Cheng Wang, Biogen Idec, Inc., 14 Cambridge Center, Cambridge, MA 02142, *yong-cheng.wang@biogenidec.com*

Key Words: Futility Analysis, Go/no-go Criteria, Proof-of-concept Study

For a proof-of-concept study, which usually is a phase 2 trial, the primary objective is usually not to demonstrate the treatment effect by a statistical hypothesis testing but to provide the go/no-go criteria for an internal decision making. Some statistical design strategies with sample size calculation and go/no-go criteria will be proposed for a proof-of-concept study. As an example, we will use an oncology phase 2 trial with a time to event variable as the primary endpoint to illustrate the proposed strategies.

A Response Adaptive Design with Pharmacogenomic Biomarkers for Targeted Therapies

✱Jens C. Eickhoff, University of Wisconsin-Madison, 250 WARF Building, Department of Biostatistics & Medical Informatics, 610 Walnut Street, Madison, WI 53726, *eickhoff@biostat.wisc.edu*; KyungMann Kim, University of Wisconsin-Madison; Jill M. Kolesar, University of Wisconsin; Glenn Liu, University of Wisconsin; Wei Huang, University of Wisconsin; Jason R. Gee, William S. Middleton Memorial Veterans Hospital, Madison WI

Key Words: adaptive randomization, personalized medicine, Bayesian logistic regression, clinical trials

Pharmacogenomic biomarkers are a critical component of a targeted therapy as they can be used to identify patients who are most likely to benefit from a treatment. New study designs are required which effectively evaluate both the diagnostic and the therapeutic intervention of pharmacogenomic biomarkers. We propose a Bayesian response adaptive design which utilizes individual pharmacogenomic profiles and patients' clinical outcomes, as they became available during the course of the trial, to assign the most efficient treatment to individual patients. A series of simulation studies were conducted to examine the operating characteristics of the proposed study design. The simulation studies show that the proposed Bayesian response adaptive design effectively identifies patients who benefit most from a targeted therapy and that there are substantial savings in the sample size requirements.

A Dose-Finding Design for Combination Therapy That Accounts for Both Efficacy and Safety

✱Ping Yao, University of Missouri-Columbia, 146 Middlebush Hall, Columbia, MO 65211, *pyqz7@mizzou.edu*; Nancy Flournoy, University of Missouri; Valerii V. Fedorov, GlaxoSmithKline

Key Words: Clinical trial, efficacy toxicity, phase II, optimal design, sequential design, bivariate probit model

We extend a two-stage design for dose-finding that accounts for both efficacy and safety (V. Dragalin, V. V. Fedorov and Y. Wu, Statist. Med. 2006) to deal with combination therapy. The bivariate probit model is used as a working model for the dose-response relationship. A "desirable therapy" is defined to be the "target dose" combination that achieves a particular probability of efficacy and safety. The goal is to estimate the dose that exists "closest" to the desirable therapy and so we seek an "optimal" design that minimizes the variance of the estimate of this dose. Although this optimal design a function of the unknown parameters and an explicit formula this optimal design does not exists, still it guides us to obtain efficient dose allocation procedures that are both safe and efficacious.

A Generalized Correlation Coefficient for the General Two-Treatment Crossover Design

✱Yi-Ju Chen, National Chung Cheng University, Department of Mathematics, No. 168, University Road, Min-Hsiung County, Chia-Yi, 621 Taiwan, *yjchen@math.ccu.edu.tw*; Vernon Chinchilli, Pennsylvania State College of Medicine

Key Words: Correlation coefficient, Crossover design, U-statistics

This study presents a general class of correlation coefficients as a platform in defining and estimating correlations for a general type of two-treatment crossover designs. We construct a Generalized Correlation Coefficient that includes the Pearson and Kendall correlation coefficients as two special members, and describe how the Generalized Correlation Coefficient can be accommodated to a general type of two-treatment crossover designs which include either at least two sequences or at least two treatment periods or both. The corresponding sample estimate of the correlation parameter is constructed that invokes the techniques of U-statistics. Certain details of the asymptotic theory for the correlation statistics are derived as well. The use of the Generalized Correlation Coefficient is demonstrated with real data examples.

Sample Size and Power Calculations for Comparing Two Means in Zero-Inflated Discrete Distribution with Application to MRI-Based Clinical Trials

✱Sungyoung Auh, National Institute of Neurological Disorders and Stroke, Clinical Neuroscience Program, Bethesda, MD 20892, *auhs@ninds.nih.gov*; Jonghyeon Kim, The EMMES Corporation; Francesca Bagnato, National Institute of Neurological Disorders and Stroke

Key Words: over-dispersion, excessive zero, estimating equation, sandwich estimator

Although contrast enhancing lesions (CELs) observed by using magnetic resonance imaging (MRI) are served as a primary outcome in most multiple sclerosis clinical trials, clinicians have started to investigate the role of drugs in reducing the formation and duration of the accumulated number of T1-hypointense lesions, namely black holes (BHs) in order to access most detriment parameter contributing to disease pathology. However, due to much less formation of BHs, a large portion of zero observations occurs. Sormani et al. (2001) conducted power analysis for the number of CELs assuming that it follows a negative binomial distribution and did not consider over-dispersion due to excess zero counts. In this paper, we propose estimating equation method based sample size estimation for comparing mean rates in zero-inflated Poisson (ZIP) and zero-inflated Binomial (ZIB) distribution.

Inference for Comparing Two Treatments Using Receiver Operating Characteristic Curve

✱Sibabrata Banerjee, Schering-Plough Research Institute, Kenilworth, NJ 07033, *sibabrata.banerjee@spcorp.com*; Sunil Dhar, New Jersey Institute of Technology; Farid Kianifard, Novartis Pharmaceuticals

Key Words: ROC, Area Under ROC, Wilcoxon-Mann-Whitney test

A reparameterized version of receiver operating characteristic (ROC) curve can be used as a measure of efficacy to compare two treatments based on the likelihood of observing a better response on one treatment than the other. The use of ROC curve to obtain the optimal point of discrimination is demonstrated using the proposed methodology. Accordingly, a point estimate and a confidence band for the measure of efficacy based on the ROC curve are obtained. We illustrate the proposed methodology with an example where a comparison is made with the point estimate and confidence interval obtained by using the nonparametric Wilcoxon-Mann-Whitney test.

A Simulation Comparison of Asymptotic Tests for 2x2 Tables of Outcome-Based Adaptive Randomization

✱Xuemin Gu, University of Texas Houston, M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, *xuegu@mdanderson.org*; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center

Key Words: Clinical trials, adaptive randomization, contingency table, hypothesis test, simulation

Outcome-based adaptive randomizations (AR) can assign more patients to the better treatment in comparative trials, and ease the ethical concerns associated with equal randomization; at the same time, certain levels of randomization are preserved to enable valid statistical inference. Due to the adaptation of patient assignment, the distribution of patient allocation is often unknown, and the results for the comparison of independent samples may not applicable to adaptive randomization. In the present study, simulations are systematically conducted across all parameter spaces to characterize the statistical properties of 9 hypothesis testing methods in 6 outcome-based adaptive randomization methods. Based on simulation results of statistical power, true type one error rate, and normality of test statistics, William's log-likelihood ratio test and generalized drop-the-urn are recommended.

Comparing the Statistical Performance of Continuous Sequential Monitoring to Group Sequential Methods for Evaluating Post-Marketing Drug Safety: A Simulation Study

✱Shanshan Zhao, University of Washington, Seattle, WA , *zhaoss@u.washington.edu*; Andrea J. Cook, Group Health Center for Health Studies; Jennifer Nelson, Group Health Center for Health Studies

Key Words: Drug Safety Surveillance, Group Sequential Methods, Simulation Study

Conducting observational post-marketing drug safety surveillance is important for detecting rare adverse events not identified pre-licensure. To detect events rapidly, continuous monitoring methods like sequential probability ratio testing (SPRT) have been proposed. However, such frequent monitoring may not be optimal or feasible, and group sequential interim analyses may be preferred. We conducted a simulation study to compare the overall type 1 error rate, power, and time-to-correct-detection of a safety problem between continuous SPRT methods and several standard group sequential monitoring approaches used in clinical trials. We vary both the group sequential testing frequency and shape of the stopping boundary. We show that power can be increased using a group sequential interim monitoring approach without substantively sacrificing timeliness to detection of an elevated safety risk.

Primary Economic Issues That Affect GDP

Les Yen, University of Phoenix/NVA; ✱Rene Wells, University of Phoenix, 11730 Plaza America Drive, Suite 200, Reston, VA 20190, *rwells7497@yahoo.com*; Eric Cramer, University of Phoenix; Angie Kruse, University of Phoenix

Key Words: GDP, Housing market, Consumer Price Index, Hypothesis tests, Mitigation plans, Currency exchange rate

The Gross Domestic Product (GDP) is one of the most comprehensive and closely watched indicator of economic activity used by the Congress, the Federal Reserve and the private sector for production, investment and employment planning. Recent man-made economic crisis including the housing market have major impact on the nation's economy. This paper examines the effect of housing market, subprime mortgages and offshoring services on GDP growth trend. We evaluated the historical data of these economic crisis and determine their effects on the GDP. Based on statistical models and hypothesis testing, we developed mitigation plans to minimize the GDP downtrend.

Effects of Alcohol and Drug Use on Human Capital: A Review

Richard Bryant, Missouri University of Science and Technology; ✱V. A. Samaranayake, Missouri University of Science and Technology, 212 Rolla Building, Rolla, MO 65409, *vsam@mst.edu*

Key Words: Alcohol/Drug Use, Human Capital, Longitudinal Data, NLSY

There has been considerable research on the effects of substance use, abuse, and dependence on measures of labor productivity such as wages, hours worked, and employment status. While not as extensive, studies on the impact of substance use on the accumulation of human capital have been the target of many studies. The results of these studies are somewhat mixed, with some showing no association between substance use and human capital measures while others show significant negative associations. However, it is not always clear that these negative associations can be interpreted as proof of the hypothesis that substance use, abuse, or dependence hinders the accumulation of human capital. In this paper, we survey some of the major research work done in this area, discuss the strengths and weaknesses in these studies, and look at the policy implications of the reported results.

A Comparison of Methods To Determine Bioequivalence of Topical Dermatological Drug Products

✱Ashlyn Hutchinson, Colorado School of Mines, Department of Mathematical and Computer Sciences, Colorado School of Mines, Golden, CO 80401, *ashutchi@mines.edu*; William Navidi, Colorado School of Mines; Annette Bunge, Colorado School of Mines; Berthe N'Dri-Stempfer, Colorado School of Mines

Key Words: Bioequivalence, Topical Drugs, Tape-stripping, Efficiency

The process for establishing bioequivalence with topical dermatological drug products is regulated by the FDA. Different formulations of dermatological drugs are applied to various application sites on test subjects. The current protocol requires eight testing sites per drug formulation, and drug levels are measured at eight distinct times over a twenty-four hour period. This is problematic, as the increased number of test sites introduces more measurement error into the experiment and a large number of subjects are required to establish bioequivalence. A new technique, the Two Time, or TT, method

requires only two application sites per subject. Simulation results indicate that the TT method is an effective method for determining bioequivalence in topical dermatological drugs. Benefits over the traditional method include the reduced time and cost due to fewer subjects and testing times.

An Estimation Method and the Appropriate Sampling Point of the Half-Life for a One-Compartment Model of a Single Bolus Intravenous Injection by a Single Sampling

✱Takashi Funatogawa, Chugai Pharmaceutical Co. LTD., 1-1 Nihonbashi-Muromachi 2-Chome, Chuo-Ku, Tokyo, 103-8324 Japan, *funatogawatks@chugai-pharm.co.jp*; Ikuko Funatogawa, Teikyo University School of Medicine; Akifumi Yafune, Clinic Sendagaya

Key Words: half-life, likelihood function, pharmacokinetics, single sampling

In clinical trials, sometimes only a single drug concentration can be measured from a patient because of the patient's burden. In this case, the sampling point is usually identical for all patients. From a single concentration, we cannot generally obtain point-estimates of each pharmacokinetic (PK) parameter. We propose a method to estimate the half-life of a one-compartment model of a single bolus intravenous injection from a single concentration at a sampling point of or after three half-lives (Funatogawa et al. 2007 J Biopharm Stat). We analytically show that the later the sampling point is the better estimate we can get. This approach is illustrated by simulated concentration data. We further propose a graphical approach which shows the shapes of the likelihood function, and confirm the appropriate sampling point for the other PK parameters, such as CL and Vd.

One Connection of State Space Model and Penalized Spline

✱Bin Zhu, The University of Michigan, 1420 Washington Heights, Department of Biostatistics, Ann Arbor, MI 48109-2029, *bzhu@umich.edu*; Jeremy M.G. Taylor, The University of Michigan; Peter Song, The University of Michigan

Key Words: State Space Mode, local linear trend model, Natural cubic spline

The state space model (SSM) is a structural model for time series data. The model assumes that the series of continuous observations are determined by unobserved latent states with some observation errors. The states themselves develop over time according to a stochastic process, for example an autoregressive process, with evolution errors. This poster is aimed to demonstrate one connection of state space model and penalized spline. The poster is organized as follows.First section introduces the state space model,especially the local linear trend model.Next section shows equivalence between local linear trend model(LLTM) and Natural cubic spline(NCS).The LLTM can be cast into linear mixed model, which is in the third section with one simulation example.We put conclusion and discussion in the last section to close the poster.

Observation-Driven State-Space Model for Categorical Time Series

✱Xiaoa Zhen, The University of Georgia, Department of Statistics, Athens, GA 30605, *zxa3wu@yahoo.com*; Ishwar Basawa, The University of Georiga

Key Words: categorical data, state-space model

We will present a new time series model for categorical data, where the response variable Yt is a random vector instead of single value. Generalized linear model for Yt conditional on past data and possible covariate have been developed in the regression approach, but it couldn't define the distribution of state parameter. An alternative approach is to model the joint distribution p(y1,., yn) from an underlying 'state' process pt using the 'state-space' formulation. Because there isn't any close form the likelihood function, it heavily relies on the complicated computing technique. We present a combination of the two approaches. More specifically, we use observation-driven state processes where the conditional distribution of pt is specified in term of the past data on Yt . We also will demonstrate how the model performs on DNA data analysis, when compared with regression type model.

Clustering Time Series Using Wavelet

✱You Jin Kim, Seoul National University, 24-206Ho College of Natural Science, Sillim 9 Dong Kwanakgu Seoul, Seoul, 151-747 Republic of Korea, *dbwls0805@hanmail.net*; Bo Ram Kim, Seoul National University; Jung Hyun Lee, Seoul National University; Sinsup Cho, Seoul National University

Key Words: spectral analysis, wavelet analysis, chaotic map approach, financial time series, clustering

In this paper a wavelet based clustering method for time series is proposed and compared with other approaches using spectral density and cross-correlations. The wavelet method focuses on the feature of time series by organizing only the detail parts of interest while the frequency domain approach uses the spectral density after removing the trend first and compares the cyclical movement only. Therefore, the wavelet method performs better for clustering time series regardless the existence of the cyclical movement. The proposed method is also compared with the chaotic map approach which uses cross-correlations of spins to measure the distance between the time series in the time domain. Since the chaotic map approach by Basalto et al. considers returns only, the wavelet method outperforms the chaotic map approach for the general financial time series.

Linex-Unbiased Filtering of Time Series

Khalil Shafie, University of Northen Colorado; ✱Saad Alkahtani, University of Northern Colorado, Applied Statistics and Research Methods, 532 McKee Hall Campus Box 124, Greeley, CO 80639, *alka8615@blue.unco.edu*

Key Words: Linex-unbiased, Moving average, Time series, M-estimate filtering, loss function

The asymmetric linex function, is used to measure the loss where negative bias and positive bias of the same magnitude have different importance. In this work, the linex unbiased filtering is introduced to smooth a time series when the error is measured in terms of linex loss. It is shown that in this case, though the filtering is not linear it is linex unbiased. Also we show that this filtering works better than any other moving average filter if the error is measured using linex function.

Time Series Models with Asymmetric Laplace Innovations

✱Alexandre Trindade, Texas Tech University, *alex.trindade@ttu.edu*; Yun Zhu, CitiGroup

Key Words: ARMA, GARCH, stable distribution, conditional maximum likelihood, financial data

We investigate properties of autoregressive moving average (ARMA) and generalized autoregressive conditional heteroscedastic (GARCH) models driven by Asymmetric Laplace (AL) noise, and discuss conditional maximum likelihood-based inference. The AL distribution plays, in the geometric-stable class, the same role played by the normal in the alpha-stable class, and has shown promise in the modeling of certain types of financial data. In the case of ARMA models, the marginal distribution of the process is derived and shown to be a mixture of AL densities. This allows for the modeling of data that is skewed and leptokurtic. Our examples show that in certain cases, ARMA models with AL noise provide a competitive fit to GARCH models with Gaussian innovations. This delivers an important message to the practitioner since the former are linear models, whereas the latter are nonlinear.

Time Series Analysis of the Fatal Accident Reporting System

✱Brian Close, University at Albany, School of Public Health, One University Place, Room 131, Rensselaer, NY 12144-3456, *brian.close@gmail.com*; Igor G. Zurbenko, University at Albany

Key Words: Times Series, Nonparametric, KZA, Kolmogorov Zurbenko, Adaptive Filter

Human driving habits strongly depend on a seasonal, weekly, diurnal structure as well as an overall trend related to a complex set of economic, meteorologic, and other factors. To see this structure requires methods for separating the different frequency bands. As a first step, this involves the separation of scales into: a long-term trend; seasonality; weekly scale; and a diurnal scale. The long term trend revealed by the data is an increasing rate of fatalities from 1996 to 2006, with jumps in 1996, and 2001. The seasonal peak for fatalities occurs in the summers and weekly peaks are on Friday and Saturday evenings, with Friday evening being the most dangerous night of the week. All results have been derived using the KZA software package from the open source statistical project R-Project.

Negative Seasonality and the Reduction of Dips in the Spectrum of a Seasonally Adjusted Time Series

✱Tucker S. McElroy, U.S. Census Bureau, 4700 Silver Hill Road, Washington, D.C., DC 20233-9100, *tucker.s.mcelroy@census.gov*

Key Words: ARIMA, Seasonality, Signal Extraction, Wiener-Kolmogorov

MSE optimal signal extraction filters produce an estimate whose dynamics differ from that of the target. The use of such filters, although widespread, is observed to produce dips in the spectrum of the seasonal adjustments of seasonal time series. These spectral troughs tend to correspond to negative autocorrelations at lags 12 and 24 in practice, a phenomenon that will be called "negative seasonality." Filters originally introduced by Wecker for stationary bi-infinite time series data ensure that the signal estimate shares the same stochastic dynamics as the original signal, and thus the problem of spectral dips is removed. This paper provides empirical documentation of negative seasonality, and extends the Wecker formulas to handle both nonstationarity and finite samples. We then apply these filters to produce seasonal adjustments without spectral troughs.

280 Business and Economics Statistics Section Speaker with Lunch (fee event)

Business and Economics Statistics Section

Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Epidemics in a Globalized World: Economic and Financial Lessons from HIV/AIDS

✱Tapen Sinha, Instituto Technológico Autónomo De México, Rio Hondo #1, Tizapan San Angel, Mexico DF, International 01000 Mexico, *tapen@itam.mx*

Key Words: HIV, AIDS, TRIPS, Africa, Economics, WTO

HIV/AIDS has gone from being the hottest topic in public policy circles to an almost completely forgotten disease. Yet, every year in the United States, 40,000 new cases are reported. The racial composition of the disease has changed radically in the United States, from being a prominent disease among whites to becoming a predominant disease among African Americans. Drugs to treat HIV/AIDS patients are available, but they are not reaching the patients due to their high cost. We will discuss the why and how of the economics and finance of HIV/AIDS. In the process, we will discuss the role of international organizations, international donor agencies, the World Trade Organization, and the controversy surrounding intellectual property rights of the pharmaceutical companies. We will draw lessons for other potential future pandemics, such as the avian flu.

281 Section on Bayesian Statistical Science Roundtables with Lunch (fee event)

Section on Bayesian Statistical Science

Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Bayesian Adaptive Trials

✱Scott M. Berry, Berry Consultants, 3145 Chaco Canyon Drive, College Station, TX 77845, *scott@berryconsultants.com*

Key Words: Bayesian, Adaptive design

In this roundtable, we will discuss the hot topic of adaptive trial design, focusing on the use of Bayesian adaptive trials. Conversation will include regulatory interactions, computing, adaptive dose finding, sample size, and dose/device selection. The leader has designed many adaptive trials for pharmaceuticals and devices and will reflect upon these examples during the discussion.

Salary, Sponsorship, and Strategy: A Case for Statistics in Sports

✱C. Shane Reese, Brigham Young University, Department of Statistics, 230 TMCB, Provo, UT 84602, *reese@stat.byu.edu*

Key Words: sports, Bayesian methods, strategy

Statistical analysis plays an important role in a variety of sports applications. In this luncheon, we will discuss not only how statisticians can have an increasing role in player performance rating, but in salary negotiations, sponsorship, and strategy. A variety of sports applications will be considered, from baseball to basketball, NASCAR to volleyball. Among the most useful tools for addressing the serious questions of salary, sponsorship, and strategy are Bayesian hierarchical models. We will discuss how hierarchical models can be effectively employed to address these important issues in sports.

282 Biopharmaceutical Section Roundtables with Lunch (fee event)

Biopharmaceutical Section
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Statistical Analysis and Practical Issues of Missing Data in Clinical Pharmacology (Phase I) Crossover Studies

✱Fang Liu, Merck Research Laboratories, 126 E. Lincoln Ave, PO BOX 2000, Rahway, NJ 07065, *fang_liu@merck.com*; Pat Larson, Merck Research Laboratories

Key Words: Missing data, Clinical Pharmacology (Phase I), Crossover, Pharmacokinetics

This discussion will be focused on the dropout and missing data issues in Clinical Pharmacology (Phase I) crossover studies, particularly in regard to pharmacokinetic data. It will identify the practical issues, examine and compare different statistical methods for handling missing data. Attendees will have the opportunity to look into various approaches used in the pharmaceutical industry and share their regulatory experiences in dealing with missing data issues in Phase I crossover studies.

Statistical Challenges and Opportunities in Oncology Clinical Trials

✱Yi He, sanofi-aventis, 1402 Fox Run Dr., Plainsboro, NJ 08536, *hydinghua@gmail.com*; Zhenming Shun, sanofi-aventis; Martin Roessner, sanofi-aventis

Key Words: Oncology, Clinical Trial, Cancer, FDA, endpoint, Design

There are more than 200 types of cancer. Due to the life-threatening nature of cancer, oncology drugs are in high demand (85 approvals so far). Recent FDA approvals provide a variety of treatment options using cytotoxic drugs, biotherapeutics, and target therapy (signal transduction inhibitors and antiagiogenic agents). Due to advances in oncology, regulatory requirements have changed and are still changing in endpoint selection, safety requirements, trial designs, etc. Additionally, oncology data are complex: tumor assessments performed with different diagnostic tools; CT scans are usually performed every 6–12 weeks; many patients have dose delays due to toxicity; patients turn to other therapies. All these issues result in missing data and potential biases that make the design and modeling of oncology clinical trials challenging. What would be our viable strategies?

Statistical Issues in Thorough QT Studies

✱Jennifer E. Hamer-Maansson, AstraZeneca Pharmaceuticals, 1800 Concord Pike, Wilmington, DE 19803, *jennifer.hamer@astrazeneca.com*

Key Words: Through QT study, ICH E14, QT prolongation, ECG

QT prolongation, thought to be a sign of increased risk for torsade de pointes (TdP), is one of the most common causes of marketed drugs being removed or restricted. The ICH-E14 guidance outlines the basic conduct of a study to determine if a new compound prolongs the QT interval. One issue of these studies is establishing a baseline measurement. Is a 10-minute, predose ECG reading adequate, or is it necessary to capture a full day of data to account for the circadian rhythm? Since thorough QT studies are commonly run using a crossover design, we would also need to discuss the requirements for a baseline measurement prior to each period. Another concern is how to analyze the data from a QT study. We'll discuss correcting the heart rate using a correction factor versus controlling for heart rate using the Holter-Bin method when analyzing the data.

Successful Partnership Between CRO and Sponsor Statisticians

✱Nfii Ndikintum, Paragon Biomedical, Inc., 9685 Research Drive, Irvine, CA 92618, *nndikintum@parabio.com*

Key Words: CRO, partnership, statistician

For various reasons (e.g., resources, expertise, etc.), pharmaceutical, biotech, and device companies are partnering with CROs for statistical services (statistics and statistical programming) support. Can sponsor statisticians count on their CRO counterparts to own their projects and proactively contribute to the success of projects? Can CRO statisticians deliver on their promise to sponsors? The purpose of this roundtable is to discuss how successful partnerships can be achieved between the CRO and sponsor statistician.

Stratified Trials with Binary Endpoints: Saving $$$ with More Efficient Analyses

✱Devan V. Mehrotra, Merck Research Laboratories, Mailstop: UG1CD-44, 351 N. Sumneytown Pike, North Wales, PA 19454, *devan_mehrotra@merck.com*

Key Words: binary data, interaction, Mantel-Haenszel test, minimum risk weights, non-inferiority, stratification

The Mantel-Haenszel test is a popular method for the analysis of stratified trials with success/failure endpoints. I will show that the M-H method has good power properties under an assumption that may be commonly violated in real trials. I will review a “minimum risk” (MR) alternative to the MH procedure (and to MH-like procedures) for both superiority and noninferiority trials and quantify power (and $$$) gains of the MR versus the MH approach.

Use of Biomarkers and Models in Exploratory Clinical Development

✱Alan Y. Chiang, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *chiangay@lilly.com*

Key Words: Biomarker, Model, Clinical Development

Over the last two decades or so, the objects and process of clinical development have been changing in response to an altered regulatory, medical, and business environment. Companies are increasingly willing to take novel compounds into man, but with the expectation of an early answer to the likely clinical and commercial success. In this roundtable discussion, I will start by giving views on how I believe appropriate use of biomarkers and models can be useful in expediting exploratory development. Then, we will focus on one or more of the following topics: the properties of biomarkers and their ‘dimensions’ (validation, innovation, proximity, specificity, and practicality); the use of biomarkers and models in decisionmaking in early drug development; and in which therapeutic areas we need good biomarkers and models.

283 Section on Health Policy Statistics Roundtable with Lunch (fee event)

Section on Health Policy Statistics
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Multiple Imputation Software for Modeling and Analysis: Reconciling the Ideal and Practical

✱Madhumita (Bonnie) Ghosh-Dastidar, RAND Corporation, 37027, *bonnieg@rand.org*

Key Words: multiple imputation, software, dimension, continuous variables, categorical variables, MI analysis

Multiple imputation has entered mainstream statistics over the past decade. Now there are several software options that allow different models for multiple imputation. Past software required joint modeling of all variables causing problems of data dimensionality and convergence, especially for categorical ontinuous outcomes. The next generation of software have adopted a conditional, sequential regression approach, implemented in IVEWare, MICE, SAS and Stata. These allow for flexible modeling and more efficient estimation. However, it is still hard to estimate many models with large datasets and small sample sizes. MI analysis routines are also limited. We will discuss specific applications of large-scale MI, limitations we face and possible ways to address these.

284 Section on Quality and Productivity Roundtables with Lunch (fee event)

Section on Quality and Productivity
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

The Many Forms of Reliability Data

✱David C. Trindade, Sun Microsystems, Inc., 6005 Assisi Court, San Jose, CA 95138-2316, *david.trindade@sun.com*

Key Words: reliability data, censoring, truncation, multi-censor, sudden access, interval censoring

Reliability data comes in many forms: exact times to failure or readout intervals, right censored, left censored, right truncated, left truncated, interval censored, double interval censored, arbitrary censoring, single censored, multicensored, random interval censoring, sudden access, and so on. How does one analyze the data properly? We will discuss the various types and consider the appropriate analysis techniques. The luncheon discussion is meant to be an overview to familiarize attendees with the forms reliability data takes and facilitate correct analysis.

Optimal Design of Split-Split-Plot and Strip-Plot Experiments

✱Bradley Jones, SAS Institute Inc., Room S4116, 100 SAS Campus Dr, Cary, NC 27513-8617, *Bradley.Jones@jmp.com*

Key Words: split-plot, optimal design, strip-plot, split-split-plot, random effects

Complete randomization for designed experiments on multistep processes may be infeasible due to logistics of the flow of units through the process. In such cases, one approach is to split batches of units coming out of the first process so the split batches see different conditions in the second process. Strip-plot and split-split-plot designs have three variances to estimate. The discussion will identify the differences in the structure of these two designs and the trade-off between efficiency and convenience in the choice of one versus the other.

285 Section on Statistical Consulting Roundtable with Lunch (fee event)

Section on Statistical Consulting
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Teaching Using Real Consulting

✱Tim Hesterberg, Google, Inc, 651 N. 34th Street, Seattle, WA 98103-8846, *timhesterberg@gmail.com*

Key Words: teaching, consulting, Practicum, hands-on

This roundtable is for people to discuss the use of real consulting projects in teaching. For example, I've taught a capstone practicum course in which teams of five students worked together for a semester to solve substantial consulting projects I'd solicited from local industry and government. I'll share some stories from those courses. I also found stories from consulting valuable when teaching more normal classes. I encourage those with other experiences using consulting in teaching to share.

286 Section on Statistical Education Roundtables with Lunch (fee event)

Section on Statistical Education
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Teaching Confounding and Multivariate Thinking in Introductory Statistics

✱Milo Schield, W.M. Keck Statistical Literacy Project, 2211 Riverside Avenue, Augsburg College, Minneaoplis, MN 55454-1350, *milo@pro-ns.net*

Key Words: Observational studies

Observational studies are common in business, economics, sociology, health, and education. Confounding is a major problem in observational studies. Confounding requires multivariate thinking—thinking on at least three variables. What should students be taught about confounding in an introductory course? Should students be taught methods of taking into account the influence of a binary confounder that allows them to work problems with numeric answers? Should students be shown that a statistically significant relationship can be made statistically insignificant (and vice versa) by a potential confounder? Should this teaching of confounders be done if it means less time to teach traditional statistical methods? If not, do we need a separate statistical literacy course that focuses more on confounding so the traditional intro course can stay focused on statistical inference?

How Is Your Biometry Course Going?

✱Dexter C. Whittinghill, Rowan University, Department of Mathematics, 201 Mullica Hill Road, Glassboro, NJ 08028, *whittinghill@rowan.edu*

Key Words: biometry, intro stat, biostatistics

If your college or university has a general, introductory statistics course (i.e., biometry or biostatistics) specifically for students in biology, how is it going? What do you cover? How do you do it? We will discuss issues such as the topics covered, the types of assessment use, what software you like, the lectures and labs, what textbook works for you and your students, and how you communicate with the "biologists over there." Whether you have taught such a course for years, or have just completed your first go at it, come and exchange ideas and syllabi. Those who are thinking of creating such a course are invited, but bring your ideas and your ears.

Myths and Fallacies in Elementary Statistics

✱Bernard Harris, University of Wisconsin, 1300 University Avenue, Statistics Dept., Madison, WI 53706, *bernard.harris@pacc-eng.com*

Key Words: statistical myths, textbook fallacies, elementary statistics

The purpose of this roundtable discussion is to highlight several myths and fallacies commonly found in elementary statistics texts. These will be augmented by anecdotal examples of scientific and technological misuse of statistical methods as a consequence of myths and fallacies. These issues will include the role of the number 30 in justifying normal distribution approximations, the treatment of outliers, pooling or nonpooling of variance estimates, calculation of quantiles, and the use of squared error as a measure of loss. The justification for the choice of various descriptive statistics is often misleading. Since I will be contradicting statements frequently found in many elementary statistics textbooks, there should be controversy and an opportunity for feedback and exchange of opinions.

287 Section on Statistical Graphics Roundtable with Lunch (fee event)

Section on Statistical Graphics, Section on Nonparametric Statistics
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Exploratory Data Analysis (EDA): Graphical Methods, Software, Applications, and Recent Developments

✱Juergen Symanzik, Utah State University, Department of Mathematics and Statistics, 3900 Old Main Hill, Logan, UT 84322-3900, *symanzik@math.usu.edu*

Key Words: Graphics, Data Mining, Visualization

At this roundtable, we will discuss interactive and dynamic statistical graphics as key concepts of EDA. We will discuss common graphical methods (such as brushing, linking, and various plot types) and freely available software (such as R, Ggobi, and Mondrian). We also will look at different applications of graphics (many from the medical field) and recent graphical developments.

288 Section on Statistics in Epidemiology Roundtables with Lunch (fee event)

Section on Statistics in Epidemiology
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Statistical Methods in Breast Cancer Research

✱Philip S. Rosenberg, National Cancer Institute, 6120 Executive Blvd, EPS 8020, Biostatistics Brach, Rockville, MD 20852, *rosenbep@mail.nih.gov*

Key Words: cancer research, statistical genetics, epidemiologic methods, breast neoplasms

A new paradigm in breast cancer research postulates that breast cancers are fundamentally several distinct clinical diseases; whether there is similar etiological heterogeneity is an active research area. In this roundtable, participants will discuss how their statistical research is exploring the clinical or etiological heterogeneity of breast cancer. What statistical methods are advancing research in this area? Are there gaps in existing methodology? With so many new developments—from molecular portraits of breast tumors and targeted therapies to novel breast cancer risk prediction models for distinct disease subtypes and populations to replicated findings from genome-wide association studies—what do statisticians need to know, with respect to both current concepts about the disease and the best statistical methods to study it?

Nonparametric Methods in Genetic Epidemiology: Multilocus Genetic Predisposition, Environmental Risk Interaction, and Complex Phenotypes

✱Knut M. Wittkowski, The Rockefeller University, Center for Clinical and Translational Science, 1230 York Ave Box 322, New York, NY 10065, *kmw@rockefeller.edu*

Key Words: u-statistics, multi-locus, genetic, epidemiology, environment, risk

As the amount of available genetic, environmental, and phenomic data increases, more complex questions arise. The questions we now face often involve several factors, rather than a single SNP or environmental risk factor. Especially for joint effect of multilocus genetic and several environmental factors on complex phenotypes, the functional models to justify linearizing transformations are often lacking, so statistical methods based on the linear model can be ill-suited. Results based on nonparametric approaches, such as u-statistics for multivariate data, depend less on assumptions that may be difficult to justify. As no single method is likely to be a panacea, the proposed roundtable is intended to provide an opportunity to discuss the pros and cons of various approaches and may lead to the proposal of a session at JSM 2009.

289 Section on Survey Research Methods Roundtable with Lunch (fee event)

Section on Survey Research Methods
Tuesday, August 5, 12:30 p.m.–1:50 p.m.

Methods and Issues in Trimming Extreme Weights in Sample Surveys

✱Frank Potter, Mathematica Policy Research, Inc., P.O. Box 2393, Princeton, NJ 08543, *FPotter@Mathematica-MPR.com*

Key Words: Sample Surveys, Weighting, Weight trimming, Estimation, Variance estimation

In survey sampling practice, analysis weights generally incorporate the initial sampling weights (the inverse of the selection probabilities) and adjustments for nonresponse and to match external totals (post-stratification or raking). Variation in the analysis weights may be by design (for example, oversampling and undersampling specific populations) or be unexpected (low response rates in subgroups). Extreme variation in these weights can result in large sampling variances and a few extreme weights can offset the precision gained from an otherwise well-designed and executed survey. Various methods are in use for developing trimming levels. The purpose of this roundtable is to discuss the current methods being used, new methods (Elliot and Little 2000, Liu et al. 2004) and issues in trimming sampling weights and the use of trimming methods in practice.

290 *Technometrics* Invited Session

Technometrics, Section on Physical and Engineering Sciences
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Technometrics Vol. 1, No. 1, 1959

✱J. Stuart Hunter, Princeton University, 157U Meadow Lakes, Hightstown, NJ 08520-3333, *stu@princeton.edu*

In retrospect the journal *Technometrics* was destined. The great war had fully demonstrated the value of statistics applied to industrial problems. By 1959, widely scattered quality control groups had coalesced into the ASQC, the ASA had formed the Section of Physical and Engineering Sciences, and the Gordon Conferences on Statistics in Chemistry and Chemical Engineering had to ration attendance. There was no technical journal devoted to research in the arts of statistics adapted to the problems of industry and the need was obvious. This brief exposition will review the events and personalities of the day that led to Vol. 1, No. 1 of *Technometrics*.

Quality Assessment for Short Oligonucleotide Microarray Data

Julia Brettschneider, University of Warwick; ✱Terence Speed, University of California, Berkeley; Francois Collin, University of California, Berkeley; Benjamin Bolstad, University of California, Berkeley

Key Words: quality control, relative log expression, microarrays, normalized unscaled standard errors, Affymetrix chips, residual scale factors

Quality of microarray gene expression data has emerged as a new research topic. Microarray quality is assessed by comparing suitable numerical summaries across microarrays, so that outliers and trends can be visualized, and poor quality arrays or variable quality sets of arrays can be identified. Since each single array comprises hundreds of thousands of measurements, the challenge is to find numerical summaries which can be used to make accurate quality calls. Several new quality measures are introduced based on probe level and probeset level information, all obtained as a byproduct of RMA/fitPLM. Quality landscapes spatially localize quality problems. Numerical chip quality measures are derived from the distributions of normalized unscaled standard errors and of relative log expressions. They are demonstrated on a variety of data sets and compared with Affymetrix' quality report.

291 Causal Inference in Vaccine Studies ●

Section on Statistics in Epidemiology, Biopharmaceutical Section, WNAR, Section on Health Policy Statistics, Biometrics Section
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Causal Inference in the Anthrax Vaccine Trials

✱Donald B. Rubin, Harvard University, Department of Statistics, 1 Oxford Street, 7th Floor, Cambridge, MA 02138, *rubin@stat.harvard.edu*

Key Words: Rubin causal model, principal stratification, direct and indirect causal effects

Randomized experiments in which humans with various vaccination regimens are challenged by lethal doses of anthrax cannot be ethically conducted. Such experiments, however, are considered ethical when conducted on guinea pigs, rabbits, or rhesus monkeys. The key to making the leap of faith that survival benefits from vaccination seen in the non-humans are relevant to humans is the use of immunogenicity measurements (e.g., antibody levels) of various kinds in all studies, obtained at many times post-vaccination. These intermediate (between treatment assignment and recording of survival outcome) measurements are used to bridge survival relationships with vaccination to humans. There is much discussion but little formal notation on how to do this, but formality explicates which assumptions are needed, and which are superfluous.

Toward Causal Inference with Interference

✱Michael Hudgens, The University of North Carolina at Chapel Hill, 3107E McGavran-Greenberg, Chapel Hill, NC , *mhudgens@bios.unc.edu*; M. Elizabeth Halloran, University of Washington

Key Words: Causal Inference, Potential Outcomes, SUTVA, Vaccines

A fundamental assumption usually made in causal inference is that of no interference between individuals (i.e., the potential outcomes of one individual are assumed to be unaffected by the treatment assignment of other individuals). However, in many settings, this assumption obviously does not hold. For example, in infectious diseases, whether one person becomes infected depends on who else in the population is vaccinated. We consider a population of groups of individuals where interference is possible between individuals within the same group. We propose estimates for direct, indirect, total and overall causal effects of treatment strategies in this setting. Using a two-stage randomization procedure, unbiased estimators of the estimates are presented. Variances of the estimators are also developed. The methodology is illustrated in two settings where interference is likely.

Vaccine Efficacy Among the Doomed When Monotonicity Is Unreasonable

✱Bryan E. Shepherd, Vanderbilt University, 1161 21st Avenue South, S2323 MCN, Nashville, TN 37232, *bryan.shepherd@vanderbilt.edu*

Key Words: principal stratification, HIV, causal effects

In vaccine efficacy trials, researchers often want to know the effect of vaccination on outcomes that are only well-defined in those who become infected. However, an analysis comparing post-infection outcomes between those infected in the vaccine and placebo arms could be biased because infection is not randomized. To assess the efficacy of a vaccine on a post-infection outcome, it may be better to focus on the principal stratum of individuals who were doomed to be infected regardless of treatment assignment. To identify who belongs to this stratum, monotonicity is often assumed, that there is no person who would be infected if randomized to vaccine but not if randomized to placebo. In light of recent results from the STEP HIV vaccine trial, such an assumption may be unreasonable. Here I discuss sensitivity analyses for relaxing monotonicity.

292 Nonignorable Missingness and Incomplete Auxiliary Data ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Hierarchical Bayes Estimation for Bivariate Binary Data with Applications to Small-Area Estimation

✱Malay Ghosh, University of Florida, Department of Statistics, Box 118545, Gainesville, FL 32611-8545, *ghoshm@stat.ufl.edu*

Key Words: Infant Mortality Rates, Gibbs Sampling, Low Birthweight, Missing Covariates

The paper addresses small-area estimation problems when the response is bivariate binary, and some of the covariates may be partially missing. Hierarchical Bayesian models which accommodate this missingness are considered. Estimators of small-area means, along with the associated posterior standard errors and posterior correlations, are provided. The method is applied to the analysis of a real data set, and the superiority of the hierarchical Bayes estimators over the direct estimators is established.

Can Calibration Be Used To Adjust for 'Nonignorable' Nonresponse?

✱Phillip S. Kott, National Agricultural Statistics Service, 3251 Old Lee Highway, Fairfax, VA 22030, *pkott@nass.usda.gov*; Ted Chang, University of Virginia

Key Words: prediction model, quasi-randomization, benchmark variable, model variable, bias, mean squared error

Calibration can be used to adjust for unit nonresponse when the model variables on which the response/nonresponse mechanism depends do not coincide with the benchmark variables in the calibration equation. As a result, model-variable values need only known for the respondents. This allows the treatment of what is usually considered nonignorable nonresponse. Although one can invoke either quasi-randomization or prediction-model-based theory to justify the calibration, both frameworks rely on unverifiable model assumptions, and both require large sample to produce nearly unbiased estimators even when those assumptions hold. We will explore these issues theoretically and with a small empirical study.

NMAR Nonresponse with Limited Covariate Information

✱Danny Pfeffermann, Hebrew University & University of Southampon, Department of Statistics,, Hebrew University, Mount Scopus, Jerusalem, 91095 Israel, *msdanny@mscc.huji.ac.il*; Anna Sikov, The Hebrew University of Jerusalem; Michael Sverchkov, BAE Systems IT/Bureau of Labor Statistics

Key Words: Imputation, Sample-distribution, Sample-complement distribution

In this talk we utilize relationships between the population, the sample and the sample-complement distributions developed previously. These relationships allow expressing the distribution of the missing target outcomes as a function of the distribution under full response and the response process. The missing outcome distribution is used for imputing the missing values and for estimating the population mean of the outcomes. The proposed approach is extended to situations where the population means of the covariates are known, but the covariates themselves are unknown for the nonresponding units. In such situations the response process for the covariates cannot be ignored even if the outcome nonresponse given the covariates is missing at random (MAR), requiring a proper imputation of the missing covariates. We illustrate the proposed approach using simulations and a real data set.

293 Communicating Technical Material with Nonstatisticians ●▲

Section on Quality and Productivity, Section on Physical and Engineering Sciences, Section on Statistical Consulting

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Extreme Makeover: Data Edition - Tables

✱Julia C. O'Neill, Merck & Co., Inc., 424 Roberts Ave., Glenside, PA 19038-3605, *julia_oneill@merck.com*

Key Words: tables, perception, scanning, digits, alignment, data analysis

Even the best educated scientists earn graduate degrees without a single course on presenting data clearly. Small changes in the formatting of tables can have a tremendous impact on the clarity of the information displayed. This presentation will cover practical guidelines for more effective communication of data in tables. A well-designed table becomes a powerful tool for data analysis. Although there are excellent publications on presenting data in tables by Ehrenberg, Few, and others, these are not easily accessible by most technologists. Their principles and guidelines have been reduced to a practical set of rules for data presenters. Before-and-after examples are used to illustrate the effect of each rule.

Perspectives on Communicating with Nonstatisticians

✱Martha Gardner, General Electric Global Research, 4031 Buckingham Dr., Schenectady, NY 12304, *gardner@research.ge.com*

Key Words: Communication, Engineers, Scientists

This talk will highlight best practices on communicating with nonstatisticians based on many years of experience conducting collaborative research with other scientists and engineers. Key topics will include becoming a team member versus being a consultant, as well as the prioritization of details. Specific examples will be shared, as well.

Extreme Makeover: Data Edition - Graphics

✱Lori Pfahler, Merck & Co., Inc., 770 Sumneytown Pike, POBox 4, WP 82-21, West Point, PA 19486, *lori_pfahler@merck.com*

Key Words: graphics, perception, communication

A well-designed graph is worth a thousand numerical analyses and an essential tool for communicating to nonstatisticians. Unfortunately, statisticians receive little training in how to effectively create graphs, and popular software for graphing data defaults to less than optimal displays. This talk will cover a set of principles and guidelines synthesized from excellent works by authorities such as Cleveland and Tufte and from years of statistical consulting experience in numerous industries. To illustrate their impact, each principle and guideline will be highlighted with interesting and relevant real-life examples. Those who apply the guidelines presented will see a dramatic improvement in the clarity of their graphics and their ability to communicate statistical analyses to nonstatisticians.

294 Modeling Long-Range Dependence and Stochastic Volatily: From Theory to Applications in Business and Economics ●▲

Business and Economics Statistics Section
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

A Bayesian Approach to Estimating the Long Memory Parameter

✱Scott Holan, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211-6100, *holans@missouri.edu*; Tucker S. McElroy, U.S. Census Bureau; Sounak Chakraborty, University of Missouri-Columbia

Key Words: Adaptive model selection, Bayesian model averaging, Hierarchical Bayes, Long-range dependence, Reversible Jump Markov Chain Monte Carlo, Spectral density

We develop a Bayesian procedure for analyzing stationary long-range dependent processes. Specifically, we consider the fractional exponential model (FEXP) to estimate the memory parameter of a stationary long-memory Gaussian time series. In particular, we propose a hierarchical Bayesian model and make it fully adaptive by imposing a prior distribution on the model order. Therefore, through an application of Bayesian model averaging, we incorporate all possible models from the FEXP class (up to a given finite order). As a result we reduce the underestimation of uncertainty at the model-selection stage as well as achieve better estimates of the long memory parameter. Additionally, we establish Bayesian consistency of the memory parameter under mild conditions on the data process. Finally the suggested procedure is investigated through simulation and the analysis of real data sets.

Inference for Long Memory Time Series with Application to Weather Derivatives Pricing

✱Nalini Ravishanker, University of Connecticut, 215, Glenbrook Road, Storrs, CT 06269, *nalini.ravishanker@uconn.edu*; Jeffrey Pai, University of Manitiba

Key Words: Long-range dependence, Temperatures data, Vector ARFIMA models, Weather derivatives

This talk describes inference for multivariate time series exhibiting long-range and short-range dependence, as well as heavy-tailed behavior, specifically for VARFIMA processes with underlying innovations assumed to be Gaussian, or Student-t or sub-Gaussian symmetric stable. Exact maximum likelihood estimation via the expectation-maximization (EM) or Monte Carlo EM (MCEM) algorithm, and fully Bayesian estimation via Markov chain Monte Carlo (MCMC) algorithms are described and compared to conditional MLEs. An application to pricing financial derivatives related to weather is discussed. This involves as a first step the modeling of multivariate daily temperatures at selected measurement sites in the US using the vector long memory models, paying attention to accommodating some degree of volatility exhibited by such series.

Independent Component Analysis for Financial Time Series

✱Daniel Peña, Universidad Carlos III de Madrid, Avda Madrid 126, Getafe, 28906 Spain, *daniel.pena@uc3m.es*

Key Words: factor models, dimesion reduction, multiple time series

Independent component analysis is presented as a way to obtain nonlinear factors in a set of time series. Several algorithms for obtaining the components are compared and it is shown the close relationship between factor models for time series and the second order blind identification algorithm proposed by Belouchrani et al. (1997). Three procedures for obtaining independent components are applied to a set of stock return financial time series. It is shown that a modification of the second order blind identification algorithm is able to show some interesting features in the data that were not reveal by standard factor analysis methods.

An Empirical Comparison of Some Parameter Estimation Methods in Stochastic Volatility Models

✱Bovas Abraham, University of Waterloo, Dept. of Statistics and Act. Sci, 200 University Ave. W., Waterloo, ON N2L 3G1 Canada, *babraham@uwaterloo.ca*; Ji Eun Choi, University of Waterloo

Key Words: Stochastic Volatility, Simulated Maximum Likelihood, Markov Chain Monte Carlo

Financial time series often exhibit time-dependent variances (volatility clustering) and excess kurtosis in the marginal distributions. One class of models which captures those features is the stochastic volatility (SV) models. In these models, time-dependent variances are assumed to be random variables generated by an underlying latent stochastic process. A standard SV model assumes that the conditional distribution of observations is normal and the volatility sequence evolves as an autoregressive sequence with log normal marginals. Exact maximum likelihood estimation is difficult in the SV models and several approximate methods are proposed in the literature. In this paper we study the Simulated Maximum Likelihood (SML) and Markov Chain Monte Carlo (MCMC) methods for estimating the parameters of a standard SV model.

295 Analysis of Medical Cost Data: Collaboration Between Health Economists and Statisticians ●▲

Section on Health Policy Statistics, WNAR
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Joint Modeling Longitudinal Semi-Continuous Data and Survival, with Application to Longitudinal Medical Cost Data

✱Lei Liu, University of Virginia, 3181 Hospital West, Charlottesville, VA 22911, *liulei@virginia.edu*

Key Words: Proportional hazards model, Frailty model, Dependent censoring, Informative drop-out, Survival analysis

It has been increasingly common to analyze simultaneously repeated measures and survival. In this talk we propose a joint model when repeated measures are semi-continuous, characterized by the presence of a large portion of zero values, as well as right skewness of positive values, e.g., monthly medical costs, car insurance annual claims, etc. A random effects two-part model is used to model respectively the odds of being positive and the actual amount of positive values. The random effects from the two-part model are then incorporated in the hazard of the failure time to form the joint model. Estimation is proceeded by Gaussian quadrature techniques conveniently implemented in SAS Proc NLMIXED. Our model is applied to monthly medical costs of 1455 chronic heart failure patients from the clinical data repository (CDR) at the University of Virginia.

The Anatomy of Health Care Cost Distributions

✱John Mullahy, University of Wisconsin-Madison, 787 WARF, 610 Walnut Street, Madison, WI 53726, *jmullahy@wisc.edu*

Key Words: healthcare costs, pareto tails, goodness of fit

This paper examines estimation of and inference about parameters of health care cost distributions. Using data from the Medical Expenditure Panel Survey (MEPS) and the Nationwide Inpatient Sample (NIS), the paper considers a set of parametric models for these distributions. The paper focuses on implications of "heavy tail" properties for estimation of and inference about parameters (e.g., (conditional-)mean costs) that play central roles in economic evaluation, risk adjustment, etc. Should health care cost distributions manifest such "Pareto-type" properties, estimates obtained from standard methods of estimation that rely on moment-based estimators (e.g., GLM) or on estimators that enforce finite moment assumptions (e.g., gamma distributions) may provide misleading guidance when such estimates are used to inform evaluations. Goodness of fit tests are deployed to assess tail fit.

Estimating Health Care Treatment Costs with Censoring and End-of-Life

✱Willard G. Manning, The University of Chicago, Harris School of Public Policy, Room 193, 1155 E. 60th St., Chicago, IL 60637, *w-manning@uchicago.edu*; Anirban Basu, The University of Chicago

Key Words: healthcare costs, skewness, censoring

Although most analyses of health-care costs use annual or per member per month methods, many applications need to examine either the lifetime costs of a diagnosis or the costs of an episode of health care or illness. These applications require robust methods to properly characterize the time-profile of costs over this episode or lifetime of the patient. Such cases are common for the analysis of clinical trials data and cost-effectiveness analysis. Standard methods do not address the analytical issues of right-censoring due to the end of the period of observation occurring before death or the end of the episode. Additionally, concerns about time to death and its implications for cost profiles have not been incorporated in traditional estimators. We have concerns about the robustness of the existing methods, especially given the extreme skewness typical of these data.

A New Semiparametric Model for Correlated Semicontinuous Data

✱Andrew Zhou, University of Washington, Seattle, WA 98195, *azhou@u.washington.edu*; Hua Z. Lin, Sichuan University

Key Words: Semi-parametric models, Health Care Costs, Skewed Data, Transformation

In this talk, we introduce a new semiparametric transformation model to fit correlated semicontinuous data. In our model, we allow the transformation to be nonparametric. Fitting the proposed model faces computational challenges due to intractable numerical integrations. We obtain the estimates for the parameter and the transformation function based on estimating equations and an approximate likelihood, which has high order accuracy but with less computational burden. We also propose an estimator for the expected value of the semicontinuous outcome on the original-scale.

296 Large Random Matrices, High-Dimensional Inference, and Future SAMSI Programs

Statistical and Applied Mathematical Sciences Institute, Section on Nonparametric Statistics, IMS
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Covariance Estimation in High-Dimensional Graphical Models

✱Bala Rajaratnam, Stanford University, Department of Statistics, Sequoia Hall, 390 Serra Mall, Stanford, CA 94305-4065, *brajarat@stanford.edu*

Recent advances in high dimensional inference have demonstrated the efficacy of Graphical models in dimensionality reduction with applications ranging from Genomics and Finance to Climate modeling. In this talk we investigate new inferential and model selection procedures for covariance estimation best suited to high dimensional graphical models. We compare and contrast the performance of our methods to existing methods. In the process, we are also led to explicitly consider the eigenvalue behavior of large random matrices.

Permutation-Invariant Regularization of Large Covariance Matrices

✱Elizaveta Levina, The University of Michigan, 1085 S. University, Ann Arbor, MI 48109, *elevina@umich.edu*

Key Words: covariance estimation, large p small n, sparsity, regularization, manifold projections, high-dimensional

Estimation of covariance matrices has a number of applications (e.g., principal component analysis, discriminant analysis, and inferring independence relationships) and the sample covariance matrix performs poorly in high dimensions. Recent research mostly focused on regularization when variables have a natural ordering. When no such ordering exists, regularization must be invariant under variable permutations. We discuss several new permutation-invariant sparse estimators, with convergence rates that make explicit the trade-offs between dimension, sample size, and sparsity of the true model. We also discuss an alternative method for finding a "good" ordering of the variables based on the Isomap, a manifold projection method. Based on joint papers with P. Bickel, A. Rothman, A. Wagaman, and J. Zhu.

On the Interplay Between Probability Distributions, Graphs, and Matrices

✱Nancy Wermuth, Chalmers/Gothenburg University, Mathematical Statistics, Sweden, *wermuth@math.chalmers.se*

The independence structure of a joint probability distribution may be captured by a graph. Such graphs consist of nodes for random variables and of edges that couple node pairs. They differ by the type and number of edges present. Questions concerning independence structures arise in particular when different studies are to be compared. Examples are to decide whether two different types of graph capture the same structure or how a given structure changes for a subset of variables and for given levels of some variables. One road to answer such questions is to represent different types of edge by corresponding binary matrices and to develop operators for these binary matrices that lead to repeated transformations of graphs. Thereby, the independence structure in an associated derived probability distribution becomes largely a consequence of the properties of the matrix operators.

297 Methodological Advances in Astronomy and Astrophysics ●

Section on Bayesian Statistical Science

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Empirical Comparisons of Computer Models for Stellar Evolution

✱David A. van Dyk, University of California, Irvine, 2206 Bren Hall, Department of Statistics, University of California, Irvine, CA 92697-1250, *dvd@ics.uci.edu*; Steven DeGennar, University of Texas at Austin; Ted von Hippel, University of Texas at Austin; William Jeffery, University of Vermont; Nathan Stein, University of Texas at Austin; Elizabeth Jeffery, University of Texas at Austin

Key Words: Astrostatistics, Bayesian Methods, Markov chain Monte Carlo, Mixture Models, Model Comparison, Dynamic Computation

Color Magnitude Diagrams (CMDs) are plots that compare stellar absolute magnitudes in different colors. High nonlinear correlations among the mass, color and surface temperature of newly formed stars induce a long narrow curved point cloud in a CMD known as the main sequence. Aging stars form new CMD groups of red giants and white dwarfs. The physical processes that govern this evolution are studied with complex computer models used to predict the plotted magnitudes as a function of parameters of scientific interest such as stellar age, mass, and metalicity. Here, we describe how we use the computer models as complex likelihood functions in a Bayesian analysis that requires sophisticated computing, corrects for contamination of field stars in the data, accounts for complications caused by binary stars, and aims to compare competing physics-based computer models of stellar evolution.

Multilevel Bayesian Modeling for Astronomical Surveys

✱Thomas Loredo, Cornell University, Space Sciences Bldg., Ithaca, NY 14853-6801, *loredo@astro.cornell.edu*

Key Words: Astronomy, Bayesian statistics, multilevel models, hierarchical models, shrinkage

Astronomical surveys arise in nearly every astronomical subdiscipline, and the resulting source or event catalogs span a huge range of sizes, from dozens to hundreds of millions. Despite this enormous diversity, many common themes arise in survey data analysis, particularly the need to account for selection effects (truncation and censoring) and various sources of measurement error. I will provide an overview of how these problems arise and are conventionally treated in diverse contexts, from surveys of ~100 minor planets to surveys of millions of distant galaxies. Of historical note, I will describe an early attempt by Eddington to handle some of these issues, resulting in a type of shrinkage estimator—in 1913. Finally, I'll describe work in progress on devising modern Bayesian approaches to handling truncation and measurement error via multilevel models for surveys.

Streaming Motion in Dwarf Spheroidal Galaxies

✱Michael Woodroofe, The University of Michigan, Department of Statisics, Ann Arbor, MI 48109, *michaelw@umich.edu*

Key Words: Empirical processes, Isotonic estimation, Threshold models

The dwarf spheroidal galaxies that surround the Milky Way are subject to gravitational forces from it that affect the velocities with which stars move within the dwarf spheroidals. The talk will describe efforts to model and estimate the effect on velocities. The primary assumption is that the effect increases with increasing distance from the center of the dwarf spheroidal; and a primary question is the extent to which the effect can be described by a threshold model. Statistical tools include the use isotonic estimation and empirical processes. Complications include missing data, selection effects and foreground stars.

298 Clustering and Sample-Size Issues for Rank Tests ●

ENAR, Section on Nonparametric Statistics, SSC, WNAR, IMS, Biometrics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Two Sample Tests for the Nonparametric Behrens-Fisher Problem with Clustered Data

✱Denis Larocque, HEC Montreal, 3000, chemin de la Cote-Sainte-Catherine, Montreal, QC H3T 2A7 Canada, *denis.larocque@hec.ca*; Riina Haataja, Tampere School of Public Health; Jaakko Nevalainen, Tampere School of Public Health; Hannu Oja, Tampere School of Public Health

Key Words: Mann-Whitney, Wilcoxon, Rank test, Cluster correlated, Behrens-Fisher

There has been a lot of interest towards generalizing rank based methods to settings involving cluster correlated data in the last few years. This talk will focus on the two sample testing problem with such data and the goal of this work is to propose Mann-Whitney type tests which are valid under minimal assumptions on the underlying distributions and clustering mechanism. A class of weighted test statistics will be introduced and compared to the recent procedures of Datta and Satten (*Journal of the American Statistical Association*, 2005) and Rosner, Glynn and Lee (Biometrics, 2006).

A Signed-Rank Test for Clustered Data

✱Somnath Datta, University of Louisville, Dept of Bioinformatics & Biostat, 555 S. Floyd Street, Louisville, KY 40202, *somnath.datta@louisville.edu*; Glenn Satten, Centers for Disease Control and Prevention

Key Words: Dependent data, Paired comparison, Sign test, Repeated measures

We consider the problem of comparing two outcome measures when the pairs are clustered. Using the general principle of within cluster resampling, we obtain a novel signed rank test for clustered paired data. We show by a simple informative cluster size simulation model that only our test maintains the correct size under a null hypothesis of marginal symmetry compared to four other existing signed rank tests; further, our test has adequate power when cluster size is non-informative. In general, cluster size is informative if the common distribution the difference in each pair within a cluster depends on the cluster size. An application of our method to testing radiation toxicity trend is presented.

Power and Sample Size Estimation for the Wilcoxon Rank Sum and Signed Rank Test

✱Bernard Rosner, Channing Laboratory, Harvard Medical School, 181 Longwood Avenue, Boston, MA 02115, *stbar@channing.harvard.edu*; Robert Glynn, Brigham & Women's Hospital

Key Words: Mann-Whitney U Test, probit transformation, ROC curve

The Wilcoxon Mann-Whitney (WMW) U Test is commonly used in nonparametric two group comparisons when the normality of the underlying distribution is questionable. The type II error of the WMW test has only been considered for specific distributions (Lehmann, 1998; Pettitt and Siskind, 1981). In this paper, we present an approach for estimating type II error of the WMW test which is applicable to any continuous distribution and also extend the approach to handle grouped continuous data allowing for ties. These results are also applied to estimating type II error for the Wilcoxon signed rank test. These results are based on SAS-callable functions to evaluate the bivariate normal integral and are easily implemented with standard software.

299 The Measure of Mollie Orshansky ●▲

Memorial, Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section, Scientific and Public Affairs Advisory Committee, Committee on ASA Archives and Historical Materials

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

The Measure of Mollie Orshansky

✱Juanita Tamayo Lott, U.S. Census Bureau, Human Resources Division, 20233, *juanita.t.lott@census.gov*; ✱Connie Citro, Committee on National Statistics, 500 Fifth Street, NW, Washington, DC 20001, *ccitro@nas.edu*; ✱Gordon Fisher, U.S. Department of Health and Human Services, 20201, *gordon.fisher@hhs.gov*; ✱John Iceland, University of Maryland, College Park, MD 20742-1315, *jiceland@umd.edu*

Key Words: poverty thresholds, Mollie Orshansky, Social Security Administration

This session will honor Mollie Orshansky, who developed the federal measure of poverty. In July 1963 and January 1965, Ms. Orshansky, a statistician for the Social Security Administration published two articles presenting a set of poverty thresholds for identifying families and persons in poverty, and analyzing the poverty population identified by these thresholds. Her poverty thresholds were adopted by the Office of Economic Opportunity on a quasi-official basis in May 1965 then designated as the Federal Government's official statistical definition of poverty by the Bureau of the Budget in 1969. The multigenerational panel will focus on Ms. Orshansky's pioneering efforts in methodology, public policy, federal and other poverty programs, and her leadership and mentoring skills that inspired younger generations in research on underserved populations.

300 Approaches to Protecting Privacy and Confidentiality in Data ▲

Section on Government Statistics, Section on Survey Research Methods

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Providing Access to Business Microdata: The International Experience

✱Stefan Bender, Institute for Employment Research, Regensburger Str. 104, Nuremberg, International 90556 Germany, *stefan.bender@iab.de*

Key Words: data access, research data centers, remote access, multiple imputation, metadata, harmonization

All statistical agencies face the dilemma of how to provide microdata access to scientists while at the same time guaranteeing the confidentiality of data providers. Traditional access modalities have included public use files, licensing, and increasingly, Research Data Centers (RDCs). However, these approaches have come under increasing criticism. As a result, new innovations have recently emerged – notably the use of multiple imputation techniques to create public use files, and technological solutions to provide remote access. This paper describes the lessons learned from the German experience, from an international workshop that was held in Nurnberg Germany and from a project proposal for the 7. EU Framework Programme. It summarizes the current best practices, together with the implications for current and future research access to microdata.

Defining Business Data Needs

✱Nicholas Greenia, Internal Revenue Service, *nicholas.h.greenia@irs.gov*

Key Words: business, data, researchers, private, public, partnership

A national tax agency's data can be leveraged through proper identification, capture, processing, and integration for statistical research. Access and budgetary constraints can be transcended with a partnership of analysts. Important questions include what fosters entrepreneurship; why the nonprofit sector's increasingly significant; outsourcing and independent contractors; the effect of true effective tax rates on economic performance. Barriers to research include public and private researchers are inadequately aware of the roles each need to play; budgets and legal authorizations limit public analysts on access and data quality options; private researchers often do not know or appreciate these constraints and the vast data potential, including its limitations. Proposed measures might serve the research community's needs outside the historical zero sum game of access and barriers.

Providing Remote Access to Business Micro Data: Lessons Learned

✱Timothy M. Mulcahy, National Opinion Research Center, 4350 East West Highway Suite 800, Bethesda, MD 20814, *Mulcahy-Tim@norc.uchicago.edu*; Pascal Heus, Open Data Foundation; Chet Bowie, National Opinion Research Center

Key Words: confidential, business-level microdata, remote access

The NORC Data Enclave, developed by NORC in conjunction with the National Institute of Standards and Technology, provides remote and on-site researcher access to confidential micro data. It combines elements from the computing and social sciences, resulting in a secure remote data access platform and high-level technical security protocols certified by NIST. Other protocols include: on-site researcher training; a review process and legal agreements to ensure that only authorized researchers from approved institutions access data; audit logs and audit trails to monitor research behavior during data access; and full disclosure review of statistical results before they are permitted to leave the secure environment. This paper summarizes some of the practical lessons learned over the past year and identifies next steps that can be applied to similar approaches.

Microdata Access

✱Felix Ritchie, Office for National Statistics, Cardiff Road, Newport, International NP10 8XG United Kingdom, *felix.ritchie@ons.gov.uk*

Key Words: Microdata, Access, Confidentiality, Disclosure

Research Data Centers (RDCs) have been the poor relations of data access. Recently, however, developments in technology have altered the landscape. Whilst effective anonymization becomes more difficult, RDCs are now able to surmount many of their drawbacks and make more use of their advantages. Hence, there has been a resurgence in RDCs and analysis of the most confidential microdata. However, this has raised new issues in terms of disclosure control of outputs, transparency of access, international data sharing and so on. This paper reviews developments and provides an overview of providing access to confidential microdata, concentrating on developments in RDCs and "data enclaves." It looks at the data access "spectrum," reviews what we know about managing RDCs, highlights known problems and unresolved issues, and discusses some of the practical barriers to improved data access.

301 Intelligent Brain Statistics ●▲

Biometrics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

A Genetic Algorithm To Choose Data Preprocessing and Kernel Ridge Regression Parameters for Optimal Prediction with fMRI

✱Rajan Patel, Google, Inc., 147 Easy St, Mountain View, CA 94043, *rajan@alumni.rice.edu*

Key Words: fmri, prediction, ridge regression, genetic algorithms

The past two Pittsburgh Brain Activity Interpretation Competitions have focused on the ability of functional magnetic resonance imaging (fMRI) data to predict dynamic experiences. Top performing teams from 2006 and 2007 used various machine learning and regression techniques from relevance vector machine (RVM) regression, kernel ridge regression, and support vector machines (SVM) to Gaussian process regression and neural networks. Judging by the success of various methods in the competition, it became clear that the choice of the method used (SVM, RVM, etc.) may not be as critical as the choices that are made in data preprocessing and model parameter selection. In this talk, I describe a genetic algorithm to help select optimal temporal filtering and spatial smoothing parameters as well as the regularization parameter for the ridge regression.

Discrimination of Multichannel Brain Signals: The Slex-Shrinkage Method

✱Hernando Ombao, Brown University, 121 South Main Street, 7th floor, Center for Statistical Sciences, Providence, RI 02912, *ombao@stat.brown.edu*

Key Words: Spectral Matrix, Periodogram Matrix, Shrinkage Estimators, SLEX, Classification

This talk will address two challenges associated with analyzing multichannel EEG data. The first is that most EEGs are nonstationary and we propose to use the SLEX which have the ability to capture localized features. The second is that these data often exhibit strong cross-correlation between channels. Consequently, the spectral matrix estimates are close to singular and inverting them may give imprecise results thereby producing poor classification rates under criteria such as the Kullback-Leibler discrimination information. We propose to estimate the spectral matrix by shrinking the slightly smoothed periodogram matrix to the identity matrix. The shrunken spectral has excellent frequency resolution and is superior to the standard periodogram smoother in terms of the squared-error risk and correct classification rates. (Joint work with Boehm and von Sachs, UCL, Belgium)

Statistical Inference for Minimum BD Estimators and Classifiers with Diverging Number of Parameters

✱Chunming Zhang, University of Wisconsin, Department of Statistics, 1300 University Avenue, Madison, WI 53706, *cmzhang@stat.wisc.edu*

Key Words: hemodynamic response function, high dimensional problems, brain fMRI, exponential family, spatio-temporal data, hypothesis test

Statistical models with diverging number of parameters arise commonly from high-dimensional regression and classification problems. We investigate the asymptotic properties, when the number of parameters tend to infinity, of the minimum BD estimator under Bregman divergence (BD), which encompasses nearly all of the commonly used loss functions in the regression, classification procedures and machine learning literature. The estimators are evaluated via simulation study and applied to real brain fMRI data and gene expression data.

Unified Framework for Modeling Binary Objects and Its Application to Amygdala Shape Analysis

✱Moo K. Chung, University of Wisconsin-Madison, 1500 Highland Avenue #437, Madison, WI 53719, *mkchung@wisc.edu*

Key Words: spherical harmonics, binary objects, brain imaging, Laplace equation

We present a novel computational framework for analyzing a population of 3D binary objects. 3D binary objects topologically equivalent to a sphere are represented as polygonal meshes and flattened to the sphere to obtain a reference coordinate system. Tracing the geodesic path of heat diffusion equilibrium, we obtain the surface flattening. The resulting flattening map enables us to parameterize the contour of the binary objects by two parameters (Euler angles) associated with the sphere. Then, instead of analyzing the binary objects in a 3D Euclidean space, we perform shape analysis along a unit sphere using spherical harmonics. Each amygdala is characterized by 768 shape parameters. This new framework is applied to modeling amygdala shape difference between 24 autistic and 23 control subjects.

Logistic Regression with fMRI Time Series Predictors

✱Martin Lindner, Columbia University, 1255 Amsterdam Ave. 10th flr MC 4409, New York, NY 10027, *martin@stat.columbia.edu*; Ian W. McKeague, Columbia University

Key Words: fMRI, brain, logistic regression, functional data analysis

We have developed a new type of functional logistic regression model involving trajectories as predictors of binary responses. The key idea is the inclusion of a time point as a parameter of interest in the model, allowing one to determine sensitive time points in the trajectories in terms of their relationship to the response. In this talk we discuss an application of the model to fMRI data, where time series from individual subjects are used as predictors to find the portion of the time course that is most predictive of the response (e.g., case or control). This allows us to identify key time intervals, associated with a certain task and brain region, which can be used to distinguish between responses. Our hope is that the approach will be useful as a diagnostic testing procedure for detecting disease. We illustrate the method using an fMRI study of a single induction of state anxiety.

302 Presentation and Interpretation of Safety Data in Clinical Trials ●▲

Biopharmaceutical Section, Biometrics Section, Committee on Applied Statisticians

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Interactive and Exploratory Safety Data Analysis and Review

✱Michael O'Connell, Insightful Corporation, 5317 Highgate Dr, Suite 213, Durham, NC 27713, *moconnell@insightful.com*

Key Words: web20, graphlets, clinical, review

Rigorous assessment of drug safety, pre- and post- drug approval, is essential to protect and promote public health. However, clinical trial designs and analysis plans are typically focused on efficacy endpoints. Indeed, while large amounts of safety data are collected in clinical trials, basic information such as which types of patients have adverse events or elevated lab values, is not well captured or summarized in statistical analyses or clinical study reports. This talk outlines a modern web 2.0 approach to interactive statistics, graphics and tabular review of clinical trial safety data. Statistical methods include hierarchical models and exploratory analyses such as machine learning and stabilized regression. Results from these analyses are presented as interactive S-PLUS GraphletsTM, ReviewletsTM, TabletsTM and GrablesTM as part of the Insightful Clinical Review solution.

Detecting Potential Safety Issues in Clinical Trials by Bayesian Screening

✱A. Lawrence Gould, Merck Research Laboratories, UGD1-88, 351 North Sumneytown Pike, North Wales, PA 19454, *goulda@merck.com*

Key Words: adverse event, odds ratio

Patients in large clinical trials report many different adverse events, most not identified in the protocol. Evaluating safety by hypothesis tests for between-group differences in AE incidence may be questionable because the data determine the hypotheses tested and lack of significance does not mean lack of risk. The Bayesian screening approach described here does not test hypotheses, is self-adjusting for multiplicity, provides a direct assessment of the likelihood of no material drug-event association, and quantifies the strength of the observed association. The approach directly incorporates clinical judgment by having the criteria for treatment association determined by the investigator(s). Application of the method to published data reveals an ability for discovering infrequently occurring potential safety issues that were not detected by other approaches.

Analysis and Summarization of Longitudinal Clinical Lab Data Toward Detection of Hepatoxicity

✱Jonathan Schildcrout, Vanderbilt University, *jonathan.schildcrout@vanderbilt.edu*

Key Words: longitudinal data analysis, pharmaceutical safety, mixture models

Towards detection of hepatic injury, standard longitudinal analyses can be used to capture mean differences in enzyme concentrations as a complex function of time. However, in most studies we do not observe a uniform shift in subject specific profiles for those on versus those off treatment. Instead subject specific responsiveness can arise in a number of different functional forms and so mean shifts may not address the relevant important scientific question. Latent mixture model can be used to estimate the probability of being a 'responder' or 'nonresponder,' and then can estimate mean shift between treatment arms within responsiveness category. Other subject specific responsiveness functional forms are also possible, and we will describe a number of approaches to summarizing longitudinal clinical lab data effectively.

Exploring the Individual

✱Matthew Austin, Amgen, Inc, MS 24-2-C, One Amgen Center Drive, Thousand Oaks, CA 91320, *maustin@amgen.com*

Key Words: patient profiles, safety analysis

In most instances the evaluation of n=1 is not of great interest. However, in safety analyses the evaluation of individuals is quite often of great concern. Many products and applications provide the ability to examine information on the individual level in clinical research and several of these will be discussed. In addition, fundamentals of patient profile construction will be explained for those interested in evaluating available applications or considering implementing their own.

303 Theil-Sen Estimates in Modern Regression ●▲

Section on Nonparametric Statistics

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Review of the Theil-Sen Estimator Plus Results on Estimating Explanatory Power

✱Rand Wilcox, University of Southern California, Dept of Psychology, Los Angeles, CA 90089-1061, *rwilcox@usc.edu*

Key Words: robust methods, explanatory power

There are two goals. The first is to review various aspects of the Theil-Sen regression estimator. The second is to consider the problem of estimating a robust version of explanatory power. It is found that in terms of mean squared error and bias, the Theil-Sen estimator performs better in simulations than the Coakley-Hettmanspserger estimator. Comments on related methods will be made.

Asymptotics of the Theil-Sen Estimator

✱Xueqin Wang, Sun Yat-Sen University, 135 Xin Gang W. Rd, Guangzhou, 510275 China, *hawkingwang@gmail.com*

Key Words: Theil-sen estimator, linear regression, U-statistics, super-efficiency

A well-known robust way to estimate the slope parameter in simple linear semiparametric regression models is to evaluate the median of all slopes of lines through pairwise sample points. The properties of the slope median, usually called the Theil-Sen estimator, have not yet been investigated completely. This talk will discuss them: Unbiasedness, consistency, asymptotic distribution and super-efficiency under mild conditions. Several examples and simulations are provided to confirm our results.

The Multivariate Spatial U-Quantile and Its Application in Theil-Sen Estimator

✱Weihua Zhou, The University of North Carolina, Charlotte, 9201 University City Blvd, Department of Mathematics and Statistics, Charlotte, NC 28223, *wzhou2@uncc.edu*

Key Words: Theil-Sen, multiple regression, multivariate regression, asymptotic, spatial U-quantile

In this paper, we use the multivariate spatial U-quantile to study the Theil-Sen estimator for the multiple linear regression model and the multivariate linear regression model. The asymptotic properties of the related estimators are explored. We also propose the corresponding algorithm and conduct the simulation to investigate the relative efficiency. This research extends the classical Theil-Sen estimator in simple linear regression into the general case.

The Theil-Sen Estimators in Semiparametric Mixed Models

Xin Dang, University of Mississippi; Xueqin Wang, Zhongshan University; ✱Hanxiang Peng, University of Mississippi, Hume Hall 305, PO Box 1848, University,, MS 38677, *mmpeng@olemiss.edu*

Key Words: Influence Function, Robustness, Semi-Parametric Mixed Model, Spatial Median, Subsampling, Theil-Sen Estimator

We propose to construct the robust Theil-Sen estimators in general regression models based on multivariate medians, to study their theoretical behaviors, and to explore their practical applications.

304 Statistical Computing with Massive Data

Section on Statistical Computing, Section on Statisticians in Defense and National Security, Section on Physical and Engineering Sciences
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

BigmemoRy: Massive Data, Shared Memory, and Parallel Computing in R

✱John W. Emerson, Yale University, 438 Humphrey St, New Haven, CT 06511, *john.emerson@yale.edu*; Michael Kane, Yale University

Key Words: R, memory, parallel, C, data, data mining

The new bigmemoRy package (described earlier by Michael Kane) allows users to store massive matrices in shared memory in UNIX environments. This has several exciting benefits. First, multiple R sessions on the same machine can share access to the same massive data set, avoiding the need for each session to manage its own copy. Second, it opens the door for development and implementation of parallel algorithms with massive data sets. This package is designed for a wide range of users, ranging from serious C and Fortran programmers to casual R users with no experience with other programming languages.

High-Performance Processing of Large Data Sets via Memory Mapping: A Case Study in R And C++

✱Daniel Adler, Georg-August University of Göttingen, Platz der Göttinger Sieben 5, Institute for Statistics and Econometrics, Göttingen, International 37085 Germany, *dadler@uni-goettingen.de*; Jens Oehlschlägel, Research Consultant; Oleg Nenadic, Georg-August University of Göttingen; Walter Zuccini, Georg-August University of Göttingen

Key Words: large dataset processing, C++, R, memory-mapping

We present the current status of a package (called 'ff') for processing large data sets that don't fit in memory. While database systems are effective for selecting subsets of complex-structured data, mass data processing in scientific contexts works on flat structures (such as vectors and matrices) whose simplicity can be exploited to enhance performance. For example mirroring regions of persistent storage into main memory (memory mapping) enables processing of the dataset in a transparent manner. We illustrate the above concepts with new R container types that mimic R vectors and matrices. In effect, these enable one to work on large data sets using familiar functions. The C++ framework allows one to specify new data types. Space-saving virtual storage modes, such as 1-bit logical or single-precision reals, are implemented.

Analysis of Huge Data Sets in S-Plus

✱Stephen Kaluzny, Insightful Corporation, 1700 Westlake N, #500, Seattle, WA 98109, *spk@insightful.com*; Michael Sannella, Insightful Corporation

Key Words: package, "large data," S-PLUS, bigdata

The S language was originally designed to analyze data that fits into memory. The bigdata package for S-PLUS introduces a system for working with arbitrarily large data sets. We describe the design of the bigdata objects and its chunking mechanism for data access. The basic S classes and methods for bigdata objects will be presented as well as some of graphical and statistical modeling methods that use bigdata. We show how to extend the methods and discuss the limitations that we have encountered.

BigmemoRy: Handling Massive Data in R

✱Michael Kane, Yale University, 25 Avalon Drive Apt 2424, Milford, CT 06460, *michael.kane@yale.edu*; John W. Emerson, Yale University

Key Words: R, data, memory, C

Multi-gigabyte data sets challenge and frustrate R users even on well-equipped hardware. C programming can provide memory efficiency and speed improvements, but is cumbersome for interactive data analysis and lacks the flexibility and power of R's rich statistical programming environment. The new package, bigmemoRy, bridges this gap, implementing persistent massive objects in memory (managed in R but implemented in C) and supporting the manipulation and exploration of these objects. This talk will show bigmemoRy in action with data from the Netflix Prize competition. This talk should be followed by the talk by John Emerson.

305 Innovations in Spatio-Temporal Cluster Detection for Infectious Disease Surveillance ●

Section on Statisticians in Defense and National Security, Section on Nonparametric Statistics, Section on Statistics and the Environment
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Prospective Clustering Analysis of Spatial-Temporal Data

✱Daniel Zeng, The University of Arizona, MIS Dept, McClelland Hall 430K, Univ. of Arizona, Tucson, AZ 85721, *zeng@email.arizona.edu*

Key Words: spatial-temporal data analysis, supper vector machines

This talk focuses on a new spatio-temporal data analysis approach aimed at discovering abnormal spatio-temporal clustering patterns. We propose a quantitative evaluation framework and compare our approach against a widely-used space-time scan statistic-based method under this framework. Our approach is based on a robust clustering engine using support vector machines and incorporates ideas from existing online surveillance methods to track incremental changes over time. Initial experimental results using both simulated and real-world data sets indicate that our approach is able to detect abnormal areas with irregular shapes more accurately than the space-time scan statistic-based method.

A Nonparametric Scan Statistic for Multivariate Spatial Biosurveillance

✱Daniel B. Neill, Carnegie Mellon University, Heinz School of Public Policy, 5000 Forbes Avenue, Pittsburgh, PA 15213, *neill@cs.cmu.edu*; Jeff Lingwall, Carnegie Mellon University

Key Words: cluster detection, event detection, spatial scan statistics, biosurveillance, nonparametric statistics

We present a nonparametric scan statistic (NPSS) for multivariate cluster detection. Given multiple streams of spatial time series data, we detect spatial regions where recent counts for some subset of streams are higher than expected. We first compute “empirical p-values” (EPV) for each stream and location, comparing current counts to the distribution of past counts. We scan over regions consisting of some subset of streams for some spatial area, and find regions with lower than expected EPV. We evaluated NPSS on disease outbreaks injected into hospital and pharmacy data from western PA. NPSS outperforms other approaches on datasets where parametric models are incorrect, and improves detection power by integrating information from multiple streams. Other advantages include characterization of affected streams, low false alarm rates, and ability to incorporate nontraditional data types.

Discrimination and Thresholding in Evaluation of Disease Cluster Detection Algorithms

✱Linda Moniz, Johns Hopkins University, 11100 Johns Hopkins Road, Applied Physics Laboratory, Laurel, MD 20723, *Linda.Moniz@jhuapl.edu*; Howard S. Burkom, Johns Hopkins University; Jian Xing, Centers for Disease Control and Prevention

Key Words: scan statistics, cluster detection

A number of researchers have evaluated disease cluster detection algorithms based upon measurements of specificity and sensitivity, usually represented in ROC or AMOC curves. However, the ROC and AMOC curves are quite sensitive to the thresholds used in the clustering algorithms. The clustering algorithms, in turn, are sensitive to both cluster geometry and density of the background (baseline) distributions. In this study we vary both cluster geometry (via injected disease cases) and thresholds (via both with selection of geographic regions according to background density and with explicit thresholding for the ROC curves). We show the evolution of the ROC curves as a function of these components and indicate when the ROC curves can faithfully evaluate the algorithms as well as when they are a misrepresentation of the results of the algorithm.

Using the Repeated Two-Sample Rank Procedure for Detecting Anomalies in Space and Time

✱Ronald D. Fricker, Jr., Naval Postgraduate School, Operations Research Department, 1411 Cunningham Road, Monteey, CA 93943, *rdfricker@nps.edu*

Key Words: biosurveillance, statistical process control, nonparametric, spatio-temporal

The Repeated Two-Sample Rank (RTR) procedure is a nonparametric statistical process control methodology that applies to both univariate and multivariate data. The method transforms a sample of data into univariate statistics; changes in the distribution of the data are then detected using nonparametric rank tests. In this discussion we explore its use as a spatio-temporal event detection and monitoring methodology for use in applications such as biosurveillance, crime mapping, or IED incidence change detection. The methodology is designed to sequentially incorporate information from individual observations as they arrive into an automated systems and thus can operate on data in real-time. Upon a signal of a possible distributional change, the methodology suggests a way to graphically indicate the likely location of the distributional change.

EWMA Smoothing and Bayesian Spatial Modeling for Health Surveillance

✱Huafeng Zhou, Merck & Co., Inc., UG1D-44, P.O.Box 1000, North Wales, PA 19454, *huafeng_zhou@merck.com*; Andrew B. Lawson, University of South Carolina

Key Words: health surveillance, spatial model, exponentially weighted moving average (EWMA), Bayesian method, Markov chain Monte Carlo (MCMC)

The development of appropriate methods of analysis for surveillance of spatial health data is of growing importance for public health applications. The need for relevant and flexible surveillance methods has been stressed many times in the context of bioterrorism. In this study a novel method for the monitoring of disease maps over time in a surveillance setting is developed. The approach relies upon the use of a spatial model which is fitted to current spatial data and is smoothed with historical spatial estimates. The method of smoothing is a vector EWMA procedure. A simulation study with a range of scenarios is presented and finally a case study of monitoring infectious disease spread in the state of South Carolina is presented. Our new approach for health surveillance is easy to implement, and can substantially reduce the computational time.

306 Analytical Customer Relationship Management ●

Section on Statistics and Marketing

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Random Forests for Internet Recommendations

Michiel van Wezel, Erasmus University; Erik Miedema, Erasmus University; ✱Rob Potharst, Erasmus University, P.O.Box 1738, Econometric Institute, Rotterdam, International 3000 DR The Netherlands, *potharst@few.eur.nl*

Key Words: recommender system, random forest, decision tree, multiple targets

We consider the use of Random Forests (RF) for internet recommendation: the most interesting items (movies, books, etc) should be recommended to a user. Recommendation problems are typically characterized by high numbers of users and items, and a sparse data matrix. To generate recommendations for large numbers of items, we adapt the standard CART algorithm to deal with multiple targets. Three splitting rules are implemented: choosing the split that minimizes the loss for a random target, choosing the split that minimizes the loss for all targets, and ignoring all targets by splitting on a random value of the split-variable. Nodes in our trees may inherit predictions when data in a target dimension is too sparse to be reliable. Experiments with the well-known Jester dataset show that accuracy and speed of the RF approach compete with the top results reported in the literature.

Modeling Customer Purchases and Lifetime Value in a Noncontractual Setting with Markov Models

✱Edward C. Malthouse, Northwestern University, 1870 Campus Drive, Evanston, IL 60208, *ecm@northwestern.edu*; Shyam Gopinath, Northwestern University; Robert Blattberg, Northwestern University

Key Words: markov chains, purchasing behavior, lifetime value, customer relationship management

Understanding customer purchasing behavior is critical for the efficient allocation of marketing resources. This study focuses on the use of Markov models in estimating customer lifetime value (LTV). While others have proposed using first-order Markov Chains with states defined by recency and possibly frequency, the first-order assumption, small set of covariates, and method of estimating transition probabilities have not been questioned or studied. We present an empirical study of whether predictive accuracy is improved by (1) relaxing the first-order assumption and (2) adding covariates such as factors accounting for seasonality and trends, previous values of recency, transaction sizes as well as time-dependent covariates for future marketing activity. We use logistic regression models to estimate transition probabilities and traversing a binary tree to estimate expected future LTV.

Random Forests for Multiclass Classification: Random Multinomial Logit Applied to an ACRM Problem

✱Anita Prinzie, The University of Manchester, MBS West - Booth Street West, Manchester, M15 6PB United Kingdom, *anita.prinzie@mbs.ac.uk*; Dirk Van den Poel, Ghent University

Key Words: analytical customer relationship management, random forests, multinomial logit, feature selection, random multinomial logit, RMNL

Multinomial Logit (MNL) is considered to be the standard in multi-class classification and is commonly applied within the analytical CRM (Customer Relationship Management) domain. Unfortunately, to date, it is unable to handle huge feature spaces typical of CRM applications. Conversely, random forests, unlike MNL easily handles high-dimensional feature spaces. This paper investigates the potential of applying the random forests principles to the MNL framework. We propose the Random Multinomial Logit (RMNL) (i.e., a random forest of MNLs) and compare its performance to that of (a) MNL, (b) random forests. We illustrate the RMNL on a cross-sell CRM problem within the home-appliances industry. The results indicate a substantial increase in performance of the RMNL model to that of the MNL model with expert feature selection. (See *www.crm.UGent.be*.)

Risk Management in Customer Relationship Management of Bank

✱Katsutoshi Yada, Kansai University, 3-3-35, Yamate, Suita, Osaka, 564-8680 Japan, *yada@ipcku.kansai-u.ac.jp*; Takashi Washio, Osaka University; Hisao Nagaoka, Kansai University; Yasuharu Ukai, Kansai University

Key Words: risk management, bank run, financial crisis, bank

A bank is a financial institution to manage and provide financial products for customers, and it has strong influence on the economy. Under the development of the financial system in our recent society, the varieties of the financial crisis are increasing. The risk management for financial crises is now becoming highly important for not only a fragile bank but also competitive one, because a groundless rumor on the bank can cause it. The objective of this paper is to construct a model that predicts the bank-run of the account holders with high accuracy. For this purpose, we have collected data on the decision of bank account holders by using questionnaire. The constructed models provided us useful knowledge for the strategic planning corresponding to the financial crises. Furthermore, our models demonstrated the possibility of its wide applicability and practicality.

Data Mining in Hospital Management

✱Shusaku Tsumoto, Shimane University, 693-8501 Japan, *tsumoto@computer.org*; Shoji Hirano, Shimane University; Hidenao Abe, Shimane University

Key Words: data mining, hospital management, risk managment, hospital information system, medical informatics

It has passed about twenty years since clinical information are stored electronically as a hospital information system. Stored data includes from accounting information to laboratory data and even patient records are now stared to be accumulated. If the implementation of electronic patient records is progressed into the improvement on the efficiency of information retrieval, it may not be a dream for each patient to benefit from the personal database with all the healthcare information, from cradle to tomb. However, although the studies on electronic patient record has been progressed rapidly, reuse of the stored data has not yet been discussed in details. In this paper, data stored in hospital were analyzed. The results show several interesting results, which suggests that the reuse the reuse of stored data will give a powerful tool to support hospital management.

307 Studies of Weather and Climate

Section on Statistics and the Environment
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

A Bayesian Reconstruction of Holocene Temperature Variation in Northern Fennoscandia

✱Lasse Holmström, University of Oulu, Department of Mathematical Sciences, P.O. Box 3000, 90014, Finland, *lasse.holmstrom@oulu.fi*; Panu Erasto, National Public Health Institute; Jan Weckstrom, University of Helsinki; Marjut Nyman, University of Helsinki; Atte Korhola, University of Helsinki

Key Words: Multiproxy reconstruction, Bayesian analysis, scale space, SiZer, BSiZer, chironomids

The arctic lakes of Northern Fennoscandia are excellent natural laboratories for the study of climatic changes. Variation in the external conditions are continuously recorded in the lake sediments for example in the form of fossils of aquatic organisms. However, different proxies, such as pollen, diatoms or chironomids have yielded different temperature reconstructions even from the same sediment record. It is therefore difficult to make any generalizations on the basis of a reconstruction made from a single record. We describe our first attempt at multi-proxy reconstruction based on two chironomid fossil records that aims to combine the signals in the records into a single "consensus" reconstruction. The reconstruction uses Bayesian modeling and the salient, time scale dependent features in the reconstruction are assessed using BSiZer scale space analysis.

Latin Hypercube Sampling To Select Model Parameters for Ensemble Weather Prediction

✱Matthew Pocernich, National Center for Atmospheric Research, 3450 Mitchell Lane, Boulder, CO 80301, *pocernic@ucar.edu*; Joshua Hacker, National Center for Atmospheric Research

Key Words: latin hypercube, weather, ensembles, sampling

Ensemble weather forecasts are an attempt to sample forecast distribution and characterize uncertainty in forecasts. Several methods exist for generating ensembles, including accounting for uncertainty inherent in initial conditions and making a collection of smaller changes within a single modeling system. An alternative is to vary individual parameters, thereby introducing variability with minimal impact on the model structure. A multiparameter approach is appealing because the forecast sample is drawn from a single distribution and can be interpreted easily. We propose Latin hypercube sampling to generate random, independent samples of four key parameters. Initial results from a month of 48-h forecasts that the multiparameter approach is competitive with approaches that vary entire closures, and may offer a viable alternative to more common, yet potentially problematic, methods.

Estimating the Global Distribution of Severe Thunderstorms and Tornadoes

✱Harold E. Brooks, NOAA/National Severe Storms Laboratory, National Weather Center Rm 4362, 120 David L. Boren Blvd, Norman, OK 73069, *harold.brooks@noaa.gov*

Key Words: climate, thunderstorms, weather

Severe thunderstorms and tornadoes represent significant threats to life and property in many parts of the world. The potential for future changes in their distribution makes it important to understand their current distribution. Differences in space and time in procedures for collecting and archiving reports of events lead to problems in using such reports in a straightforward manner. As a result, relationships between environmental conditions and events have been developed so that differences in the frequency of those environmental conditions can be used as proxies for the events themselves. Problems at each step of the process will be discussed, as well as estimates of the current distribution and possible changes in the future.

Spatial Hierarchical Modeling of Weather Extremes from a Regional Climate Model

✱Dan Cooley, Colorado State University, Department of Statistics, Fort Collins, CO 80523-1877, *cooleyd@stat.colostate.edu*; Stephan Sain, National Center for Atmospheric Research

Key Words: climate change, extremes, precipitation

Regional climate models (RCMs) are tools which allow scientists to begin to understand how different forcings may affect climate. There has been some statistical work done to characterize the difference in mean behavior between control and future scenarios as predicted by RCMs. The goal of this work is to characterize the extremes as produced by a RCM and to additionally examine the difference in extremes between a control and future scenario. To characterize the spatial behavior of extremes we construct a hierarchical model. The data level is formed by the point process representation of extremes, and the process level is based on a conditional autoregressive (CAR) model since our data are on a regular lattice. To our knowledge, this is the first work which spatially models the shape parameter of the extreme value distribution.

K-Nearest Neighbor as Tool for Downscaling Information from Climate Models

Matthew Pocernich, National Center for Atmospheric Research; ✱David Yates, National Center for Atmospheric Research, 3450 Mitchell Lane, Boulder, CO 80301, *yates@ucar.edu*

Key Words: time series, downscaling, climate change, weather

Current climate models provide information on a large scale and often poorly replicate features of daily weather such as local precipitation and temperature. There are many consumers of weather data such as water utilities and power utilities who will be greatly affected by a changing climate. In order to use existing models to plan for future conditions, these decision makers would like realistic weather times series which reflect future climatic conditions at the local level. The k-Nearest Neighbor method resamples historical data in a biased manner to reflect future conditions. This method relies on first-order Markov properties to create realistically autocorrelated time series. Problems arise when changes occur which exceed the range of the historical weather data. Other modifications need to be made to accommodate climate change forecasts.

308 Student Paper Competition: Bayesian Methods in Inverse Problems, Imaging, and Engineering ●

Section on Bayesian Statistical Science

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Detecting Unexpected Residual Spectral Structure in X-Ray Astronomy

✱Jason Kramer, University of California, Irvine, CA , *jskramer@uci.edu*; David A. van Dyk, University of California, Irvine

Key Words: astronomy, multi-scale, Markov random field, MCMC, EM, Poisson

As X-ray telescopes have become more powerful, datasets collected for high-energy spectral analysis have become larger and increasingly complex. Spectral analysis is the study of the energy distribution of photons arriving from an astronomical source and gives clues as to the temperature, composition, and physical processes of the source. Photon counts are modeled as a Poisson process while blurring of energies, stochastic censoring, and background contamination are built into a highly structured hierarchical model. We use a finite mixture model with one component for a physics-based parametric spectral model and a second component that uses a smoothing multi-scale prior distribution to account for unexpected residual spectral structure in the data. Bayesian methodology and computing are used to fit the model and to compute estimates, error bars, and credible intervals.

Regularized Posteriors in Linear Ill-Posed Inverse Problems

✱Anna Simoni, Toulouse School of Economics (GREMAQ), 21, Allée de Brienne, F, Toulouse, 31000 France, *simoni.anna@gmail.com*; Jean-Pierre Florens, Toulouse School of Economics (GREMAQ - IDEI)

Key Words: Gaussian Process, Conditional Probability, Tikhonov Regularization, Sufficient Statistics, Posterior Consistency

This paper studies Bayesian solution for a signal-noise problem in which we observe a noisy transformation of the infinite dimensional parameter that we want to recover. Actually, we deal with a functional equation in Polish spaces which is ill-posed because of compactness of the operator characterizing it. We follow a Bayesian approach, hence the solution is the posterior distribution of the parameter of interest. To get rid of continuity problems caused by infinite dimension of the spaces, we compute a regularized version of the posterior distribution and we guess it is solution of the inverse problem. Frequentist consistency is proved under regularity conditions of the true parameter. Finally, we extend our analysis to the case with unknown operator and the case with different operators for each observation. Monte Carlo simulations confirm good properties of our estimator.

A Bayesian Image Analysis of Change in Tumor/Brain Contrast Uptake Induced by Radiation

✱Xiaoxi Zhang, Pfizer, Inc., 235 E 42nd Street, New York, NY 10017, *xiaoxi.zhang@pfizer.com*; Timothy D. Johnson, The University of Michigan; Roderick J. Little, The University of Michigan; Yue Cao, The University of Michigan

Key Words: hidden Markov random fields, Quantitative MRI, reversible jump MCMC, Swendsen-Wang algorithm, trans-dimensional proposal

This work is motivated by a quantitative Magnetic Resonance Imaging study of the change in tumor/brain contrast uptake induced by radiation. The results show a transient period at which contrast uptake in the tumor is maximal; suggesting an optimal time to initiate chemotherapy during the course of radiotherapy. A notable feature of the data is spatial heterogeneity. We introduce a latent layer of discrete labels and employ a Gaussian hidden Markov random field (MRF) model that respects this feature. Conditional on the hidden labels, the observed data are assumed independent and normally distributed. We estimate the MRF regularization parameter, treat the number of MRF states as a parameter and estimate it via a reversible jump Markov chain Monte Carlo algorithm with a novel and nontrivial implementation. We examine the performance of our method in a simulation study and on real data.

Analyzing Space-Time Sensor Network Data Under Suppression and Failure in Transmission

✱Gavino Puggioni, Duke University, Department of Statistical Science, NC , *gp10@stat.duke.edu*; Alan Gelfand, Duke University

Key Words: Wireless Sensor Networks, Suppression, Hierarchical Models, Spatio-temporal models, Information Loss

In this paper we present a model-based analysis of the effects of suppression and failure in data transmission with sensor networks. Sensors can be created to collect data at very high temporal resolution. However, during periods when the process is stable, transmission of such high resolution data would carry little additional information (i.e., all of the data that is collected need not be transmitted). In particular, when there is cost to transmission, we shall consider suppression. Additionally, in practice, we will experience failures in transmission. Evidently, both suppression and failure lead to information loss which will be reflected in inference. Our effort here is to assess the impact of such information loss under varying extents of suppression and failure. We consider two illustrative process presenting analyses of suppression and failure using hierarchical models.

309 NSF Programs Supporting Statistics Education and Strategies for Becoming a Successful Investigator

Section on Statistical Education, Social Statistics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

NSF Programs Supporting Statistics Education and Strategies for Becoming a Successful Investigator

✱Ginger Holmes Rowell, National Science Foundation, 4201 Wilson Blvd, Division of Undergraduate Education, Suite 835, Arlington, VA 22230, *growell@nsf.gov*; ✱Dennis K. Pearl, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210-1247, *dkp@stat.osu.edu*; ✱Joan Garfield, The University of Minnesota, 315 Burton Hall, 178 Pillsbury Drive, S.E., Minneapolis, MN 55455, *jbg@umn.edu*

Key Words: statistics education, grant opportunities, proposal writing

A number of National Science Foundation (NSF) divisions offer a variety of grant programs that support innovations in teaching and learning in statistics. These programs will be discussed along with examples of successful projects. Experienced statistics education grant award winners will also provide feedback about their proposal writing experiences.

310 Time Series Topics

Business and Economics Statistics Section
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Bartlett-Type Corrections for Score Test Statistics on Gaussian ARMA Models

✱Bernardo M. Lagos, University of Concepción, Department of Statistics, Esteban Iturra s/n, Barrio Universitario, Cocepción, International 160-C Chile, *bla@udec.cl*; Pedro A. Morettin, Institute of Marine Science, University of S„o Paulo

Key Words: ARMA models, Score statistics, Bartlett-type correction

In this article we compute the corrected score statistic according to Bartlett-type proposed by Cordeiro and Ferrari (1991), used to test the null hypothesis concerning some parameter vector of interest of an ARMA model, assumed to be Gaussian, stationary and invertible. We also consider the situations where nuisance parameters are present, useful in many practical cases. The formulas are written in matrix form, appropriate for the use of symbolic or numerical languages. The performance of the tests is assessed through an estimate of the rejection probability error and an estimate of the power, via Monte Carlo simulation study. Many cases of Bartlett-type Score test is better than the original version. Particularly cases where the parameters are close to the critical regions, not stationary model.

Models for Continuous Dynamical Processes with Bounded Support

✱Amanda R. Cangelosi, Utah State University, 3900 Old Main Hill, Logan, UT 84322, *amanda.c@aggiemail.usu.edu*; Mevin B. Hooten, Utah State University

Key Words: Bias-Correction, Differential Equations, Numerical Approximations, Population Dynamics, Truncated Normal, Hierarchical Models

Models for natural nonlinear processes, such as population dynamics, have been given much attention in applied mathematics. For example, species competition has been extensively modeled by differential equations. It is of both scientific and mathematical interest to implement such models in a statistical framework to quantify uncertainty in the presence of observations. This study offers an alternative to common ecological modeling practices by using a bias-corrected truncated normal distribution to model the observations and latent process, both having bounded support. Parameters of an underlying continuous process are characterized in a Bayesian hierarchical context, utilizing a fourth-order Runge-Kutta approximation.

A Multiscale Variance Stabilization for Binomial Sequence Proportion Estimation

✱Matthew A. Nunes, University of Bristol, School of Mathematics, University of Bristol, University Walk, Bristol, International BS8 1TW England, *Matt.Nunes@bristol.ac.uk*

Key Words: Binomial random variable, Gaussianization, Haar-fisz, sequence probability estimation, variance stabilization

There exist many different wavelet methods for classical nonparametric regression in the statistical literature. However, techniques specifically designed for binomial intensity estimation are relatively uncommon. In this article, we propose a new technique for the estimation of the proportion of a binomial process. This method, called the Haar-NN transformation, transforms the data to be approximately normal with constant variance. This reduces the binomial proportion problem to the usual `function plus normal noise' regression model and thus any wavelet denoising method can be used for the intensity estimation. We demonstrate that our methodology possesses good Gaussianization and variance-stabilizing properties through extensive simulations, comparing it to traditional transformations. We also explore the efficacy of our method in a real application.

The ARIMA (0,1,2)(0,1,1) 32 Model and Confidence Intervals

✱Sujata Mukhopadhyay, NIL, B 14/168, Applied Practitioner, Honorary, Kalyani, 741 235 India, *sujmuk@hotmail.com*

Key Words: Reformulating Synodic Time Series, Synodic ARIMA Model, Forecasts, 95% Confidence Intervals, Program in FORTRAN/BASIC, Program in C+

This ARIMA model is lunar phase–based time domain analysis (STAT' 2000). Synodic time series (53rd ISI 2001) lead to a seasonal ARIMA (0,1,2)(0,1,1) 32 Model (JMS 138, 1, 2006).The synodic scale was first studied by explanatory model (3rd bnp 2001).Then, outliers were tested with Dixon test statistics (abstract RSS 2006). Recently, the autumn equatorial forecast of 4906 was proved less than that of summer equatorial forecast 2354 due to this underlying synodic time scale (abstract ICSPRAR ISI 2008). Here, in continuation count data [Ztp] used for ARIMA forecasts [Wtp] with that of available measurement data [Xtp] and corresponding count data [Ytp], results are shown through 95% confidence intervals. Also, time series reformulation with FORTRAN and BASIC changed to C+ programs are discussed. Computations were done at CSU (CSSC) ISI Calcutta.

Simple Tests for Short Memory in ARFIRMA Models

✱Timothy Hughes, Boise State University, 305 N 17th St Apt 207, Boise, ID 83702, *thughes20@hotmail.com*

Key Words: Long Memory, Delta Method, Sample Autocorrelations, Hypothesis Testing

For an ARFIMA (p,d,q) process, it is often priori known that the process involves long memory. The objective of this thesis is to test whether the long memory time series needs other short memory parameters in the fractional model. The asymptotic distribution of sample autocorrelations is further developed in this paper. Simple functions of sample autocorrelations are considered in a way that they satisfy root-n consistency and their limiting distributions are normal regardless of the true parameter, d. The advantages of using sample autocorrelations for testing autoregressive moving-average parameters lie in ease of the computation and use of the established asymptotic theories. In addition, new graphical identification methods for short memory are presented. These proposed methods are then applied to long memory time series data in finance and geology.

Test of Normality of a Long-Memory Sequence Using the Empirical Moment Generating Function

✱Sucharita Ghosh, Swiss Federal Research Insitute WSL, Birmensdorf, CH-8903 Switzerland, *rita.ghosh@wsl.ch*

Key Words: Empirical moment generating function;, Time series;, Integral transforms;, Long range dependence;, Goodness-of-fit tests;

Several authors have considered using the empirical generating functions (emgf) for goodness-of-fit tests. In this paper, we focus on time series observations with long-range dependence and test the null hypothesis that the marginal distribution is normal using the emgf. In particular, we consider derivatives of the emgf and use these to develop asymptotic critical regions for a graphical procedure. Some simulated series as well as some global climate observations illustrate the procedure.

311 Statistical Methods for Biomedical Research ●▲

WNAR, Section on Statistical Graphics, Section on Statistics in Epidemiology, Biometrics Section
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Weighted Area Under the Receiver Operating Characteristic Curve and Its Application to Gene Selection

✱Jialiang Li, National University of Singapore, 117546 Singapore, *stalj@nus.edu.sg*; Jason P. Fine, University of Wisconsin-Madison

Key Words: Gene selection, Empirical distribution, Location-scale model, Partial area under the curve, Random threshold, Weighted area under the curve

Partial area under the ROC curve (PAUC) has been proposed for gene selection in Pepe et al. (2003) and thereafter applied in real data analysis. It was noticed from empirical studies that this measure has several key weaknesses, such as an inability to reflect nonuniform weighting of different decision thresholds, resulting in large numbers of ties. We propose the weighted area under the ROC curve (WAUC) in this paper to address the problems associated with PAUC. Our proposed measure enjoys a greater flexibility to describe the discrimination accuracy of genes. Nonparametric and parametric estimation methods are introduced, including PAUC as a special case, along with theoretical properties of the estimators. We also provide a simple variance formula, yielding a novel variance estimator for nonparametric estimation of PAUC, which has proven challenging in previous work.

A Bayesian Method for Cross-Trial Inference in the Noninferiority Setting

✱Siobhan Everson-Stewart, University of Washington, 916 NE 70TH ST APT 10, Seattle, WA 98115, *spes@u.washington.edu*; Scott S. Emerson, University of Washington

Key Words: clinical trials, non-inferiority, Bayesian methods

In noninferiority clinical trials, researchers would often like to show that the new therapy would have been proven superior to placebo, had one been included in the trial, in addition to being noninferior to the active control. Such proof requires cross-trial inference, which previously required strong, often unrealistic assumptions. For example, most traditional methods require that the constancy assumption holds (i.e., that the effect of any one therapy remains constant from trial to trial). In therapeutic settings with changing patient populations, however, this assumption is often highly suspect. We present a Bayesian method for cross-trial inference when treatment effect changes as a function of the study population. The performance of this method is examined under a variety of circumstances.

A Varying Coefficient Method for Analyzing Longitudinal Data with Nonignorable Dropout

✱Jeri E.F. Harwood, University of Colorado Denver, 4200 E. 9th Ave., Campus Box C245, Denver, CO 80262, *ForsterHarwood.Jeri@tchden.org*; Samantha MaWhinney, Health Sciences Center University of Colorado Denver ; Diane Fairclough, University of Colorado Denver

Key Words: B-splines, Clinical trials, HIV/AIDS, Longitudinal data, Nonignorable dropout, Varying-coefficient models

Dropout is common in longitudinal studies and is nonignorable when the probability of dropout depends on unobserved outcomes even after conditioning on available data. In this case, mixture models can be used to account for the relationship between a longitudinal outcome and dropout. We propose a varying-coefficient mixture model using natural cubic B-spline basis functions (VCM_ns) to semiparametrically model the outcome-dropout relationship. This method is computationally stable, highly flexible and relatively simple to implement. We apply the VCM_ns and comparable methods to an HIV/AIDS clinical trial that shows evidence of nonignorable dropout. In addition, we conduct simulation studies to evaluate performance and compare methodologies. The simulation studies suggest that the VCM_ns is an improvement over existing methods when the dropout mechanism is nonlinear.

Estimates of Information Growth Accommodating Model Misspecification

✱Abigail Shoben, University of Washington, Department of Biostatistics, Campus Mail Stop 357232, Seattle, WA 98195, *ashoben@u.washington.edu*; Kyle Rudser, University of Washington; Scott S. Emerson, University of Washington

Key Words: clinical trials, information growth, longitudinal data

In group sequential clinical trials, it is necessary to estimate the amount of information at interim analysis times relative to the information at the final analysis. This is often the ratio of sample sizes, if only one measurement is made per individual. However, as noted by Wu and Lan (1992, Biometrics, 48, 765-779), when the statistic of interest is a change over time such an approach overestimates the information. In this talk, we explore problems that can result when analysis of randomized controlled trials is based on standard linear models. Specifically, when the true pattern of dose response does not correspond exactly to the linear contrast used to measure treatment effect, information growth curves estimated with standard methods may be nonmonotonic. We give examples in which these standard estimates are inappropriate and demonstrate approaches that provide improved estimates.

Dynamic Transitional Modeling Through Bayesian Variable Selection in Disease State Prediction

✱David Engler, Brigham Young University, 230 TMCB, Department of Statistics, Provo, UT 84602, *engler@byu.edu*; Brian Healy, MGH Biostatistics Center

Key Words: Bayesian variable selection, Markov model, ordinal response

In multiple sclerosis (MS), the main clinical outcome measure is an ordinal score, the expanded disability severity score (EDSS). One of the main goals of MS research is to accurately predict the future EDSS state of patients. It is expected that the probability of transitioning from one state to another will vary across subjects. Moreover, under the assumption of a Markov process, the order of the Markov chain is often unknown and must be estimated. We propose use of Bayesian variable selection in Markov transitional models that allows for the identification and estimation of subject-specific transition probabilities as well as the estimation of the order of the Markov process. Methods are assessed using both simulated data and data collected from the Partners MS center in Boston, MA.

Using Nonlinear Mixed Models To Analyze Characteristics of Discounting Behavior

✱Susan K. Mikulich-Gilbertson, University of Colorado Denver, 4200 East 9th Avenue, #C268-35, Denver, CO 80262, *Susan.Mikulich@uchsc.edu*; Thomas J. Crowley, University of Colorado Denver; Laetitia Thompson, University of Colorado Denver

Key Words: nonlinear mixed models, delay discounting, random effects, hyperbolic function, substance users, adolescents

Prior two-stage analyses suggest adults discount smaller rewards more rapidly, real and hypothetical rewards similarly, and that substance users (SU) discount delayed rewards more rapidly than nonusers. SAS NLMIXED was used to evaluate different models for comparing discounting in adolescents and to assess reward characteristics. Multiple delay intervals and reward paradigms were evaluated in 40 SUs and 40 controls. Nonlinear models of mathematical functions were compared between groups and assessed reward type and magnitude. Exponential models found that SUs discounted delayed rewards more rapidly but hyperbolic functions fit best and found no differences. Including random rate effects improved model fit but hampered convergence. Multivariate hyperbolic models found both groups discounted smaller rewards more rapidly, but only SUs discounted hypothetical and real rewards differently.

Exploratory Graphics for Functional Actigraphy Data

✱Juergen Symanzik, Utah State University, Department of Mathematics and Statistics, 3900 Old Main Hill, Logan, UT 84322-3900, *symanzik@math.usu.edu*; William Shannon, Washington University School of Medicine

Key Words: Actigraphy, Functional Data Analysis, Visualization

Actigraphy is an emerging technology for measuring a patient's overall activity level nearly continuously over time. An actigraph is a watch-like device attached to the wrist or ankle that uses an accelerometer to measure movement every minute, or even more often. Actigraphy data can best be described as functional data. In this talk, we will present exploratory graphical displays for various components of functional actigraphy data.

312 Bayesian Methods for the Social Sciences ●

Social Statistics Section, Section on Bayesian Statistical Science, Section on Government Statistics
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Hierarchical Nonlinear Profile Analysis of Dichotomous Item Responses

✱Steven A. Culpepper, Metropolitan State College of Denver, Department of Human Services, Campus Box 12, P. O. Box 173362, Denver, CO 80217-3362, *sculpepp@mscd.edu*

Key Words: Hierarchical Nonlinear Models, Profile Analysis, Item Response Theory, Experiment, Map Comprehension

This study linked hierarchical nonlinear profile analysis (NPA) of dichotomous responses with a larger family of item response models. The NPA method offers several benefits over previous profile analysis methods: 1) NPA relies upon maximum likelihood rather than different dissimilarity measures that produce different results; 2) item and person parameters are computed during the same estimation step with an appropriate distribution for dichotomous variables; 3) the estimation procedure produces standard errors of profile coordinates; and 4) subject-level variables can be included to predictor profile parameters. An application examined experimental differences among 288 subjects' topographic map comprehension. The model produced a measure of overall test performance, in addition to correspondence weights with an item-difficulty profile and item-discrimination profile.

Determining the Number of Classes in Growth Mixture Modeling: Covariate, Data Generation Scheme, and Class Specificity

✱Libo Li, University of California, Los Angeles, Integrated Substance Abuse Programs, 1640 S. Sepulveda Boulevard, Suite 200, Los Angeles, CA 90025, *lilibo@ucla.edu*; Yih-ing Hser, University of California, Los Angeles

Key Words: Growth mixture model, class enumeration, data generation scheme, class specificity, covariate, adjusted BIC

Our studies on growth mixture modeling (GMM) show that the inclusion of covariates to the class-invariant GMM fitting models is generally helpful for their correct class enumeration under model consistency. Although the class-specific GMM models that are less restrictive and are capable of parameter and membership recovery perform very poorly for class enumeration, especially in small samples, our studies suggest that as long as some informative covariates can be found, these disadvantages could be overcome even with a small sample. Our extensive data generation schemes discussion also calls into question the current common practice in GMM application, which segregates class enumeration and final interpretive models by using different fitting models. Additionally, we find that the adjusted Bayesian Information Criterion is the best index for class enumeration in the GMM context.

Applying Bayesian Networks to the Examination of Student Retention and Graduation

✱Xiaohong Li, Sam Houston State University, 200 Charles Holtom 10E, College Station, TX 77840, *xxl001@shsu.edu*; Melinda M. Holt, Sam Houston State University

Key Words: conditional probability, conditional independent, Bayes net, Bayesian network, joint probability

This paper presents a method for constructing Bayesian Networks (BN) to model student success at a public university based on the prior information from Sam Houston State University's (SHSU) database of First Time Freshman cohorts from 2000 to 2006, We describe the use of BBN to examine the probabilistic relationships among student outcomes for the purpose of identifying families of factors associated with college success. We look at the affect of factors such as full or part time, instate or out of state, living on or off campus and admission types, etc. on retention and graduation rates. This will be accomplished by using Bayesian methods to analyze SHSU persistence data by using a computer application program - NeticTM3.25 (http://*www.norsys.com*).

Predicting Electoral Vote Totals for the Presidential Election

✱Christopher J. Rigdon, Arizona State University, 944 S. Terrace Rd. #204, Tempe, AZ 85281, *crigdon@asu.edu*; Steven Rigdon, Southern Illinois University Edwardsville; Edward Sewell, Southern Illinois University Edwardsville; Sheldon Jacobsen, University of Illinois-Champaign

Key Words: presidential election, polling data, Bayesian statistics

During the U.S. Presidential Race of 2008, there will be hundreds of polls conducted in each state to determine which candidate is most likely to win the election at any given time. The objective of this research is to use polling data to derive an empirical posterior distribution for the number of electoral votes each candidate will receive. By expanding on the work of Kaplan and Barnett (2003), a model is derived which considers undecided voters which can have a significant effect on election day. The model performed quite well in the 2004 election, missing the correct number of electoral votes for each candidate by only two.

Election Prediction Using the Kalman Filter

✱Steven Rigdon, Southern Illinois University Edwardsville, 1333 Science Building, Edwardsville, IL 62026, *srigdon@siue.edu*; Michael Siekerka, Southern Illinois University Edwardsville; Christopher J. Rigdon, Arizona State University

Key Words: Election polls, survey sample, Kalman filter

The Kalman filter is used to monitor the current state of the electorate in each state. This information, together with a dynamic algorithm for obtaining the posterior probability distribution of the number of electoral votes for each candidate, will be used to track the posterior probability that each candidate will win the November 2008 election.

Spatial Empirical Bayes: Borrowing Strength Through Spatial Dependence

✱Natalya Verbitsky, Mathematica Policy Research, Inc., 600 Maryland Ave, SW, Suite 550, Washington, DC 20024, *nverbitsky@mathematica-mpr.com*; Stephen W. Raudenbush, The University of Chicago

Key Words: empirical Bayes, small area estimation, Chicago neighborhood measures, collective efficacy, spatial dependence, ecometrics

Social scientists have long been interested in differences between neighborhoods in their social structure and informal organization. In this paper we show that precision and validity of ecometric measures can be improved by exploiting spatial dependence of neighborhood processes within the framework of empirical Bayes shrinkage. We compare three estimators of a neighborhood social process: ordinary least squares (OLS), empirical Bayes estimator based on the independence assumption (EBE), and empirical Bayes estimator that exploits spatial dependence (EBS). Under our model assumptions, EBS dominates EBE and OLS in estimating the true latent social process. A cross-validation study using 1995 and 2002 data from the Project on Human Development in Chicago Neighborhoods is conducted and shows that that the empirical benefits of EBS approximate those expected under our model assumptions.

Rating Crop Insurance Contracts with Skewed Distributions: A Bayesian Approach

✱Vitor Ozaki, University of S„o Paulo, av. padua dias, 11 agronomia, piracicaba, 13418900 Brazil, *vitorozaki@yahoo.com.br*; Ralph Silva, Campinas State University

Key Words: crop insurance, Bayesian hierarchical model, premium rate, Skew-normal distribution, spatial correlation

Unlike the US, which developed its crop insurance program long ago, Brazil started in 2004. Given the increasing interest in crop insurance, accurate calculation of the premium rate is of great importance. This study addresses the crop-yield distribution issue and its implications to pricing the contracts. The analysis models the dynamic structure of the data and the spatial correlation through Bayesian Hierarchical models. We choose a more flexible probability density function known as the Skew-Normal distribution estimating all the parameters through the MCMC algorithm. Comparing the premium rates derived in our model (BR) with the empirical rates (ER) method, commonly used by the insurance companies, the ER are higher than BR for low-risk areas and lower in high-risk areas. It means insurers are underpricing the insurance contract in high-risk areas and overpricing in low-risk areas.

313 Improving Survey Quality: Mode Effects, Coverage, and Forms ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Examining Mode Differences in the Health Information National Trends Survey

✱Jeffrey Kerwin, Westat, 1650 Research Blvd., Rockville, MD 20850, *jeffreykerwin@westat.com*; David Cantor, Westat

Key Words: survey methods, mode effects, response effects

The Health Information National Trends Survey (HINTS) provides health communications researchers with ongoing U.S. population estimates on cancer-related knowledge, attitudes, and behavior. Due to decreasing telephone survey response rates, the most recent HINTS is a mixed-mode data collection with dual sampling frames: an RDD telephone survey combined with a mail questionnaire sent to a sample of addresses from a national listing provided by the U.S. Postal Service. Data collection from over 7,000 households was completed in early 2008. There are numerous open-ended questions in the telephone mode where interviewers classify responses into fixed categories. The mail survey provides these categories to respondents in a closed-ended format. This presentation will report on the extent of mode differences found in the responses to these HINTS measures.

An Analysis of Mode Effects in Three Mixed-Mode Surveys of Veteran and Military Populations

✱Boris Rachev, ICF International, 9300 Lee Highway, Fairfax, VA 22031, *brachev@icfi.com*; Michael Yang, ICF International; Diana Davis, ICF International; Ronald Szoc, ICF International

Key Words: mixed-mode, mode effects, web survey, mail survey, response rate

Studies on mixed-mode survey designs and survey mode effects so far have produced mixed findings. We hope to contribute to the literature by examining mode effects between mail administration of paper surveys and Web administration of their electronic equivalents in three recent mixed-mode surveys of veteran and military populations. The three surveys are: US Army MWR Leisure Needs Survey, Survey of Veteran Satisfaction with the VA Home Loan Guaranty Process, and 2008 Veterans Burial Benefits Survey. Initial study results indicate that the mail survey mode is preferred over the Web, with some notable variation in mode preference across surveys. The analysis examines how selected survey elements such as response rates, respondent profiles, item response rates, and key statistical estimates differ between mail and web. T- tests and ANOVA are used to compare modes.

Using Multiple Data Sources To Identify Types and Sources of Coverage Errors on an American Indian Reservation

✱Laurie Schwede, U.S. Census Bureau, 15016 Butterchurn Lane, Silver Spring, MD 20905, *Laurel.K.Schwede@census.gov*

Key Words: census, coverage, American Indian reservation, 2006 Census Test, methodology, living situations

This paper reports results of a study to improve understanding of household coverage on the Cheyenne River Indian Reservation in the 2006 Census Test. Matched data sources from 30 reservation households are used to identify and explore types and sources of coverage error to suggest improvements to enumeration methods, not to estimate coverage rates. Matched data include: 1)the actual Update/Enumerate household roster; 2) the Census Coverage Measurement (CCM) household roster collected 3 months later; 3) the CCM interview audiotape; and 4) an immediate ad-hoc debriefing to resolve any coverage anomalies. We identify types of coverage errors (e.g., omissions, erroneous enumerations); types of persons missed; and sources of error (e.g. mobility, living situations, cultural differences, non-city addresses, questions, recall issues). Suggestions are made to improve reservation enumeration.

Using Administrative Records To Improve Within-Household Coverage in the 2008 Census Dress Rehearsal

✱Timothy L. Kennel, U.S. Census Bureau, DID HQ-6H106B, 4600 Silver Hill Road, Washington, DC 20233, *Timothy.L.Kennel@census.gov*; Dean Resnick, U.S. Census Bureau

Key Words: Administrative Records, Logistic Regression, Undercoverage, Census

The paper presents a method using administrative records and models to increase the accuracy and efficiency of the 2008 Census Dress Rehearsal coverage follow-up operation. As a part of the 2008 Census Dress Rehearsal, some households will receive a follow-up interview. One of the goals of the coverage follow-up operation is to find and correct within household undercoverage. Since within household undercoverage is rare, a simple random sample of households is not efficient at targeting households with initial undercoverage. This paper presents one method used to target households that may have incorrectly enumerated the number of people in the household. Using models built from Census 2000 and administrative records, the probability that a household is missing one or more individuals in the initial 2008 Census Dress Rehearsal is calculated.

2007 Census Bilingual Forms Design Test

✱Kelly Govern, U.S. Census Bureau, 4600 Silver Hill Road, DSSD HQ-4H264B, Washington, DC 20233, *Kelly.A.Govern@census.gov*

Key Words: Bilingual, Census, Stratification, Spanish, Response

The 2007 National Census Test provided an opportunity to study the response results of two Spanish/English bilingual questionnaire designs in areas that contained a relatively heavy concentration of Spanish-speaking people with a need for English language assistance. Results from the 2005 National Census Test showed that while the bilingual form significantly increased self-response rates nationally, it yielded higher nonresponse rates for the household items compared to the English-only form. The 2007 National Census Test was conducted as a follow-up to the 2005 Test to determine if the bilingual form item nonresponse issues could be resolved through improved forms design and updated content. This paper compares the two 2007 bilingual form designs with variations in cover letter placement to determine the better design and then compares that design to the English-only form.

Impact of New Income Questions on Reported Income in the National Survey on Drug Use and Health (NSDUH)

✱Jeremy Aldworth, RTI International, *jaldworth@rti.org*; Dicy Painter, SAMHSA; Tania Robbins, RTI International; David Heller, RTI International

Key Words: Household survey, reported income, questionnaire change, imputation, Social Security, Supplemental Security Income

The NSDUH is an annual national survey of substance use and mental health measures in the U.S. civilian population aged 12 or older. We show that we were able to reduce the number of income questions in the 2006 NSDUH, thereby reducing respondent burden, without significantly impacting estimates of reported income. In this survey, 5.7% of the 67,802 respondents received the reduced set of questions, and analyses showed an impact in only one case. A question on Social Security income was eliminated, and as a result, it appeared to be confused with Supplemental Security Income, which was retained. A decision was made to reintroduce questions about Social Security in later surveys. Simulated imputation analyses showed that the new questions would have little impact on imputed estimates, and timing data analyses showed a 24% reduction in median time to complete the reduced set of questions.

Calendarization of the Goods and Services Tax Data: Issues and Solutions

✱Martin Beaulieu, Statistics Canada, R.H.Coats Building, section 11-Q, 100, Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *martinj.beaulieu@statcan.ca*; Benoit Quenneville, Statistics Canada

Key Words: calendarization, administrative data, indicator series

Statistics Canada uses the Goods and Services Tax (GST) data as auxiliary data in several sub-annual surveys. Given that GST remittance can be done on an annual, quarterly, monthly or on a more frequent basis, the GST data must be calendarized. Calendarization means that the data have to be standardized so that they refer to a common reporting period. The calendarization system, which was built and implemented in 2002, is now being revised in order to improve the quality of the data by solving some of the issues that were raised by the users. This talk will discuss these issues and the solutions to be implemented.

314 Inference from Longitudinal Surveys and Double Sampling ●

Section on Survey Research Methods, Section on Government Statistics

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Efficiency Comparison for Constrained Generalized Least Squares Estimators in a Panel Survey

✱Jason C. Legg, Iowa State University, CSSM, *jlegg@iastate.edu*; Cindy Yu, Iowa State University

Key Words: National Resources Inventory, Covariance estimation

One approach for estimation in a longitudinal panel survey is to use marginal generalized least squares with panel totals. For indicator variables, the sum of the totals for the indicators often is constrained to equal a known total, such as the population size. One approach to satisfying the constraint is to use the same weighting matrix for each marginal regression. An alternative approach is to use a direct estimator of the covariance matrix for the weights for each marginal regression and ratio adjusting to match the constraint. The second approach is a linear combination of estimators of the best linear unbiased estimators, but is more computational intensive than the first approach. We evaluate the loss in efficiency from using the same weight matrix compared to direct covariance estimators in the context of the National Resources Inventory design.

Matching Surveys in Longitudinal Studies

✱Cynthia B. Augustine, RTI International, 3040 Cornwallis Road, Cox 351, Research Triangle Park, NC 27709, *caugustine@rti.org*; Kristine L. Rae Olmsted, RTI International; Scott Ginder, RTI International

Key Words: longitudinal surveys, sampling precision, tobacco

This research addresses the issue of linking baseline and follow up surveys in longitudinal studies. Social Security Number (SSN) is commonly used to link survey waves, but reluctance to provide this information reduces matches. Optical scanning errors can also produce mismatches between otherwise matching surveys. Two linking methods were developed for a smoking survey that had no other identifying information except SSN. These methods were tested using a similar smoking survey on a parallel population where verified matching existed to examine the accuracy of the matching methods. One method was developed using demographic and survey variables unlikely to change between waves. A second method was developed for instances where SSNs existed but numbers were possibly transposed or poorly scanned. The cost-benefit of these matching methods will be discussed in terms of sampling precision.

Exploration of the Use of Empirical Bayes Procedures for Estimating Changes in Occupancy Rate and Persons per Household

✱Lynn Weidman, U.S. Census Bureau, 8413 Canning Terrace, Greenbelt, MD 20770, *lynn.weidman@census.gov*; Don Malec, U.S. Census Bureau

Key Words: population estimates, Empirical Bayes

The Census Bureau is carrying out research with the intent of improving the housing unit-based method for estimating population totals. One approach is to use the Decennial Census as the baseline for measures and update them annually using American Community Survey data to estimate their change since the census year. This study looks at the possibility of using an Empirical Bayes approach to produce county estimates of change with smaller variances than direct county estimates. Results from national and state models are compared and the problem of modeling states with few counties is considered. Data from the 1990 and 2000 long form samples are used to represent ACS and data from the 1990 and 2000 short form are used as independent variables in the models. In the actual application, sources for the latter would be the Master Address file and Administrative Records.

Using the SIMEX Method To Estimate Temporal Change for a High-Scoring Group

✱Andrea R. Piesse, Westat, 1650 Research Blvd., Rockville, MD 20850, *AndreaPiesse@westat.com*; Randy Sitter, Simon Fraser University; X. C. Feng, Simon Fraser University

Key Words: Regression toward the mean, SIMEX; temporal change

An important objective for longitudinal surveys is the estimation of temporal change. The simple estimation of change over time for a group that has high (or low) initial values for the variable under study produces biased estimates of true change when the variable is subject to measurement error. This bias is commonly known as the regression toward the mean artifact. We propose and outline the use of the simulation extrapolation (SIMEX) method (Cook and Stefanski, 1995) for survey estimates of change to adjust for regression toward the mean. The results of a simulation study that examines the performance in terms of bias, variance, and coverage of this application of the SIMEX method will be presented.

Efficient Calibration for Variants of Double Sampling

✱Takis Merkouris, Athens University of Economics and Business, Department of Statistics, 76 Patision Street, Athens, 10434 Greece, *merkouris@aueb.gr*

Key Words: two-phase sampling, generalized regression estimator, calibration estimator, optimal regression estimator, composite estimator

Results from recent work on integration of data from independent surveys through regression is adapted to certain variants of nested and non-nested double sampling arising in practice. A computationally simple calibration scheme that gives rise to efficient regression estimators for characteristics surveyed in the "second-phase" sample is investigated within a framework of optimal regression estimation. This "one-step" calibration scheme makes efficient use of all available auxiliary information in the first and second phase samples and greatly facilitates variance estimation. Applications to business and household surveys having the discussed double sampling designs are also considered.

Analysis of an Outcome-Dependent Enriched Sample Using Semiparametric Likelihood Method

✱Qing Kang, North Dakota State University, Fargo, ND 58105, *qing.kang@ndsu.edu*; Christopher I. Vahl, North Dakota State University

Key Words: choice-based sample, semiparametric likelihood, profile likelihood, partial likelihood, consistency

An outcome-dependent enriched sample results from adding a random sample to a stratified sample, where the stratification is based on the outcome. Consider binary outcomes with probability determined by predictors according to a given model. This study applies the profile, the partial, and the weighted likelihood methods to estimate model parameters from an enriched sample without imposing any parametric forms for the distribution of predictor variables. Although the three methods all consistently estimate model parameters, estimators given by the partial likelihood method have variance dramatically larger than the other two. Our simulation results indicate the profile, and the weighted likelihood methods perform well for moderately large sample sizes. Guidelines for the optimum allocation of sample sizes are also given for this enriched sampling scheme.

New Regression-Type Estimators in Multiphase Sampling

✱Mariyam Hafeez, University of the Punjab, Planning & Development Cell,, University of the Punjab, Lahore, International 54000 Pakistan, *umemariyam@hotmail.com*; Muhammad Q. Shahbaz, COMSATS Institute of Information Technology

Key Words: Multiphase Sampling, Regression Estimators, Ausiliary Variables, Mean Square Errors

Multiphase sampling has been widely used in surveys for estimation purpose. Several estimators have been proposed from time to time for use in two phase sampling. In this paper some new regression type estimators have been developed for use in multiphase sampling. These estimators have been developed by using single and two auxiliary variables. The mean square errors of the proposed estimators have also been derived. The empirical comparisons of the developed estimators have been made in order to pick the best estimator.

315 Infectious Disease Epidemiology ●

Section on Statistics in Epidemiology, Biometrics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Estimation of the Basic Reproduction Number for Infectious Diseases

✱Shenghai Zhang, Public Health Agency of Canada, Centre for Communicable Diseases and Infectious Control, 100 Eglantine Driveway, Rm2309, Ottawa, ON K1A 0K9 Canada, *s.zhang@rogers.com*; Ping Yan, Public Health Agency of Canada

Key Words: Infectious disease, basic reproduction number, reproductive power

It is important to estimate the basic reproduction number of an infectious disease for understanding the propagation of the infectious disease. However, the estimation procedures of the basic reproduction number are based on the data generated in the outbreak of an infectious disease. The data-generating process of the outbreak of the infectious disease has many features. The data is generated by not only a biological process but also control measures. The estimation of the basic reproduction number (or reproductive power) will be discussed by considering the aspects of the data-generating process and the model which can captures the essential characteristics of disease transmission in a real community.

Confidence Intervals for Estimates of Excess Mortality

✱Al Ozonoff, Boston University School of Public Health, 715 Albany Street, Crosstown Center, Biostatistics, 3rd floor, Boston, MA 02118, *aozonoff@bu.edu*; Xiaopeng Miao, Boston University School of Public Health; Po-Yung Cheng, Centers for Disease Control and Prevention; William W. Thompson, Centers for Disease Control and Prevention

Key Words: Influenza, Surveillance, Excess mortality, Confidence intervals

The 122 Cities Mortality Reporting System (CMRS) is one component of the national influenza surveillance program operated by the Centers for Disease Control and Prevention (CDC). Participating cities submit weekly counts of pneumonia and influenza (P&I) deaths and all-cause (AC) deaths. CDC models the square root of the arcsine of the ratio of P&I deaths to AC deaths using an Andrews robust regression method. The model estimates both a seasonal baseline and threshold value for determining influenza-associated mortality. Prior work has demonstrated the benefit of robust regression methods to reduce estimation bias. In this work we investigate variability, and consider the statistical issues involved in constructing confidence intervals (CIs) for estimates of excess mortality. Using historical data from the 122 CMRS, we compare model-based CIs to empirical CIs from a resampling procedure.

Comparison of Methods for Estimating Influenza-Associated Deaths from the 1972/1973 Through 2003/2004 Seasons

✱Hong Zhou, Centers for Disease Control and Prevention, 1600 Clifton Road,, Atlanta, GA 30333, *fwd6@cdc.gov*; William W. Thompson, Centers for Disease Control and Prevention; Po-Yung Cheng, Centers for Disease Control and Prevention; Eric Weintraub, Centers for Disease Control and Prevention ; Cecile Viboud, National Institutes of Health; Corinne Ringholz, National Institutes of Health; Joe Bresee, Centers for Disease Control and Prevention; David Shay, Centers for Disease Control and Prevention

Key Words: influenza-associated deaths, Serfling regression model, robust regression model

Several methods for estimating influenza-associated deaths were compared, including Serfling regression models, robust regression models, and a Serfling spline regression model. Weekly underlying pneumonia and influenza (P&I) deaths were modeled from the 1972/73 through 2003/04 respiratory seasons. Influenza epidemic periods were defined by using either P&I deaths or influenza deaths. Ninety-five percent confidence limits were used to define epidemic weeks. The average annual numbers of epidemic weeks estimated by these methods were similar (range11-14) and significantly correlated (r>=.44). Epidemic weeks defined by one Serfling model and the Serfling spline regression model were significantly correlated with epidemic weeks defined by influenza viral data (r>=.52). Annual influenza-associated deaths estimated by all methods were similar (range 5643-8403) and highly correlated (r>=.85).

Comparison of Seasonal Regression Models for Estimating Influenza-Associated Mortality in the US: A Simulation Study

✱Po-Yung Cheng, Centers for Disease Control and Prevention, 1600 Clifton Road, MS A32, Atlanta, GA 30333, *pcheng@cdc.gov*; William W. Thompson, Centers for Disease Control and Prevention; Cecile Viboud, National Institutes of Health; Corinne Ringholz, National Institutes of Health; Tuyen Do, Centers for Disease Control and Prevention; David K. Dhay, Centers for Disease Control and Prevention

Key Words: Serfling regression model, influenza, deaths

Based on US mortality patterns for the 1973/1974 through 2002/2003 respiratory seasons, we simulated baseline deaths and influenza-associated deaths for underlying pneumonia and influenza deaths (P&I), underlying respiratory and circulatory deaths (R&C), and all-cause deaths (AC). The seasonal average of influenza-associated deaths for P&I, R&C, AC deaths were 8,300, 30,700, 44,300, respectively. We fit three different Serfling regression models: models using 5-year baselines, models with influenza deaths as an indicator of influenza activity, and robust regression models. Estimates of annual averages of influenza-associated deaths for the three models for P&I deaths were 7,700, 7,200, 7,600, respectively; 29,000, 28,000, 27,900 for R&C deaths; and 42,900, 41,700, 40,500 for AC deaths. The simulated excess deaths and estimates from various models were highly correlated (0.94-0.99).

A Bayesian Method for Inferring Transmission Chains in a Partially Observed Epidemic

✱Jaideep Ray, Sandia National Laboratories, MS 9159, PO Box 969, Livermore, CA 94550-0969, *jairay@somnet.sandia.gov*; Youssef M. Marzouk, Sandia National Laboratories

Key Words: Bayesian inference, transmission chains, smallpox, Abakaliki

We present a Bayesian approach for estimating transmission chains and rates in the Abakaliki smallpox epidemic of 1972. The epidemic affected 30 individuals in a community of 74; only the dates of appearance of symptoms were recorded. Our model assumes stochastic transmission of the infections over a social network. Distinct binomial random graphs model intra- and inter-compound social connections, while disease transmission over each link is treated as a Poisson process. Link probabilities and rate parameters are objects of inference. Dates of infection and recovery comprise the remaining unknowns. Distributions for smallpox incubation and recovery periods are obtained from historical data. Using Markov chain Monte Carlo, we explore the joint posterior distribution of the scalar parameters and provide an expected connectivity pattern for the social graph and infection pathway.

A Discrete-Time Sequential Test for Constant Fatality Rate of an Emerging Epidemic with Applications to Severe Acute Respiratory Syndrome in Hong Kong and Beijing

✱Ying Xu, University of Hong Kong, Rm518, MengWah Complex, University of Hong Kong, Pokfulam Road, Hong Kong, International , *tinayxu@hku.hk*; S. F. Yip, University of Hong Kong; K. F. Lam, University of Hong Kong

Key Words: fatality rate, martingale difference, sequential test, severe acute respiratory syndrome (SARS)

During the course of an emerging epidemic, the fatality rate is a popular quantitative index to measure the lethality of the disease and to address the severity of the epidemic. In general, a decreasing fatality rate over time reflects the effectiveness of the clinical treatment, whereas a constant one may indicate the futility. A discrete-time sequential test, built with the martingale difference approach and using the daily aggregated data of deaths and recoveries, is proposed for testing the hypothesis of a constant fatality rate, against a decreasing one, in a timely manner throughout the outbreak of an infectious disease. Simulation studies showed that the proposed test performs well and is extremely sensitive in picking up the decreasing fatality rate over time. The severe acute respiratory syndrome (SARS) data in Hong Kong and Beijing are also used for illustration.

HWR: An Adaptive Anomaly Detection Algorithm

✱Sylvia Halasz, AT&T Labs - Research, 195 Oxford Ave, Fair Haven, NJ 07704-3133, *halasz@att.com*; Colin Goodall, AT&T Labs - Research; Arnold Lent, AT&T Labs - Research

Key Words: time series analysis, anomaly detection, simulated outbreaks

The HWR algorithm for anomaly detection in univariate time series is based on the Holt-Winters method without trend and with two kinds of multiplicative seasonality. HWR also uses run rules from statistical quality control. We will report on applications both to public health data with simulated outbreaks and to completely simulated data. Anomaly detection signals are based on different functions of the sequence of standardized residual scores. We will compare results using different ways of combining these functions, including creating one continuous score, providing several different types of alerts, and combining these approaches.

316 Teaching Advanced Topics ●

Section on Statistical Education

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

The Importance of Making the Distinction Between 'Invalid' and 'Inefficient' Procedures

✱Laurence D. Robinson, Ohio Northern University, Department of Mathematics, Ada, OH 45810, *L-Robinson.1@onu.edu*

Key Words: Invalid and inefficient procedures, multiple comparisons

In many statistics courses and textbooks, students are taught to analyze data using one particular method, often that method which has been determined to be most efficient by some criterion. Students typically come away from their courses with the false impression that methods other than that which they have learned in class are necessarily invalid, when in fact other valid methods often exist, but are inefficient compared to the methods learned in class. One particular prominent example, involving multiple comparisons, will be discussed.

The Negative Hypergeometric Distribution

✱John J. Kinney, Retired, 12 E. Caramillo St., Colorado Springs, CO 80907, *jackkinney@msn.com*

Key Words: Sampling without replacement, negative hypergeometric distribution, mean and variance by recursion, approximations, Mathematica

Sampling without replacement from a finite population leads to the negative hypergeometric random variable. The mean and variance of this distribution are found using a recursion. Implications of the mean value are discussed together with approximations by other distributions. Instructors in courses in probability are encouraged to include this distribution.

Confidence Intervals for the Difference Between Two Means

✱Weiwen Miao, Haverford College, 370 Lancaster Ave, HAVERFORD, PA 19041, *wmiao@haverford.edu*

Key Words: Behren-Fisher Problem, Coverage probability, Expected length, preliminary test, Shapiro-Wilk test, Welch-Satterthwaite

The talk compares three confidence intervals for the difference between two means when the distributions are non-normal and their variances are unknown. The confidence intervals considered are Welch-Satterthwaite interval, the adaptive interval that incorporates a preliminary test (pre-test) of symmetry; and the adaptive interval that incorporates the Shapiro-Wilks test as a pre-test. The adaptive intervals use the Welch-Satterthwaite interval if the pre-test fails to reject symmetry (or normality) for both distributions; otherwise, apply the Welch-Satterthwaite interval to the log-transformed data, then transform the interval back. Our study shows that the adaptive interval with pre-test of symmetry has the best coverage among the three intervals considered.

Using Factor Tables To Understand Autoregressive Models

✱Wayne A. Woodward, Southern Methodist University, Department of Statistical Science, Dallas, TX 75275, *waynew@smu.edu*; Henry L. Gray, Southern Methodist University; James R. Haney, Southern Methodist University; Alan C. Elliott, University of Texas Southwestern Medical Center

Key Words: time series, stationarity, autoregressive, frequency domain, time domain, decomposition

When teaching introductory classes in time series analysis, students are often left lacking when it comes to understanding the basic properties of an AR(p) model fit to a data set. Most time series texts do not provide much depth of understanding for the student (or time series practitioner) about the basic features of the model. We discuss the use of a factor table for conveying useful information about the model, including stationarity as well as time and frequency domain information. We also discuss a simple continued fraction approach for using the factors in the factor table to decompose the time series into its additive components. The factor table and decomposition can be used in a time series course to improve the understanding of the features of an AR model so that students no longer primarily view the model as a "black box" from which they can obtain forecasts, etc.

Popcorn Popping Yield: An Investigation

✱Irina A. Kukuyeva, University of California, Los Angeles, 8125 Math Sciences Bldg., Statistics Department, Box 951554, Los Angeles, CA 90095, *ikukuyeva@gmail.com*; Jean Wang, University of California, Los Angeles; Yuliya Yaglovskaya, University of California, Los Angeles

Key Words: Popcorn Yield, Experimental Design

What is a world without popcorn? People have been popping popcorn since pre-historic times. And they still do. We will introduce yet another method. What started out as a class project turned out to be a pilot study on popcorn yield using a replicated factorial design. What factors affect how much and how fast popcorn pops? The authors examine four factors (time in the microwave, oil amounts, brand of popcorn and container for the microwave) to answer this question. The purpose of our paper is to give an overview of the results and motivate further research on popcorn rates.

Using Board Games and Mathematica To Teach the Fundamentals of Finite, Stationary Markov Chains

✱Roger Bilisoly, Central Connecticut State University, 1615 Stanley St., P.O. Box 4010, New Britain, CT 06050-4010, *bilisolyr@ccsu.edu*

Key Words: Statistics education, Markov chains, Board games, Mathematica

In a stochastic processes class, Markov chains are an important example and simple board games can be used to illustrate their fundamental properties. For example, a looping board game (like Monopoly) consists of all recurrent states, and a game where players win by reaching a final square (like Chutes and Ladders) consists of all transient states except for the recurrent ending state. With the availability of symbolic mathematics packages, complex games can be analyzed, for example, the mean number of visits to the transient states and the stationary probabilities for recurrent states are both easily computed. This talk shows how to analyze simple board games with Mathematica and includes analyses of Chutes and Ladders, the Gambler's Ruin and Monopoly. For example, we demonstrate that the most likely square (besides Jail) in Monopoly is Illinois Avenue.

317 Networks

Section on Statistical Computing, Section on Statisticians in Defense and National Security, WNAR

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Methods for Time-Dependent Network Analysis

✱Andrew C. Thomas, Harvard University, One Oxford Street, 7th Floor, Cambridge, MA 02138, *acthomas@fas.harvard.edu*; Joseph Blitzstein, Harvard University

Key Words: social, networks, hypergraph

We illustrate the extension of methods for network analysis beyond a static link approach, as commonly seen in social network analysis, and a dyadic approach, as seen in most types of network analyses, to a more general dynamic-linked higher-order graph system. We demonstrate these methods on a long-term data set of dynamic social connections for the inference and prediction of social patterns.

Collapsibility for Directed Acyclic Graphs

✱Xianchao Xie, Harvard University, Department of Statistics, Science Center, 1 Oxford Street, Cambridge, MA 02215, *xxie@fas.harvard.edu*; Zhi Geng, Peking University

Key Words: Graphical model, directed acyclic graph, collapsibility, model collapsibility, marginalization

Collapsibility means that certain aspects of a model are preserved after marginalization over some variables. In this talk, we present conditions for three kinds of collapsibility for directed acyclic graphs (DAGs): estimate collapsibility, conditional independence collapsibility and model collapsibility. We show that unlike in graphical loglinear models or hierarchical loglinear models, the estimate collapsibility and the model collapsibility are not equivalent in DAG models. We discuss the relationship among them and illustrate how the results obtained can be applied in simplifying the inference problems in DAGs. Algorithms are also given to find a minimum variable set containing a subset of variables of interest onto which a DAG model is collapsible.

Combining Covariate Data and Graph Regularization in a Semisupervised Setting

✱Mark Culp, West Virginia University, *mculp@stat.wvu.edu*; George Michailidis, The University of Michigan

Key Words: semi-supervised, model, regularization, graph, convergence, semi-parametric

In this talk, we present a general framework that combines feature-based (X) data and graph-based (G) data for prediction of the response (Y). We examine a model fitting approach to fit eta=Xb+f(G) with link eta=g(u), coefficient b, and f is a function over the vertices of the graph. The procedure is semi-supervised in nature (trained on the labeled and unlabeled sets), requiring iterative algorithms for fitting this model. We discuss convergence properties of this procedure and provide a generalization to local scoring algorithms in semi-supervised learning. Applications in text data, proteomics and genomics are discussed. Specifically, in text problems the observations are papers, Y is the paper's category, X is the journal that the paper was published, and G is the co-citation network (vertices represent papers, and edges are citation counts between two vertices).

Grid Computing

✱Abdullah A. Alnoshan, The George Washington University, 344 Maple Avenue, #321, Vienna, VA 22180, *noshan1@hotmail.com*; Shmuel Rotenstreich, The George Washington University

Key Words: Grid Computing, Computational, Computer Science, Programming, Software

This paper and research is on grid computing.

Fault Detection in Embedded Networked Sensing

✱Sheela Nair, University of California, Los Angeles, *snair@ucla.edu*; Mark H. Hansen, University of California, Los Angeles

Key Words: sensor networks, data quality, fault detection, spatial temporal modeling

Recent advances in sensor technology, computing, and low-power communication have facilitated the development of embedded networked sensing (ENS). A major issue limiting the use of ENS is quality of sensor data, which can be prohibitively faulty. Fault detection systems for ENS face unique statistical and computing challenges. Data collected by ENS are often highly complex and non-stationary. Algorithms that are both computationally feasible and scalable must be designed so that data quality can be assessed in near-real time. We discuss the use of signatures for fault detection, borrowing on similar ideas used in fraud detection. The proposed algorithm adaptively estimates and tracks sensor signatures as data are collected. We also outline current work in spatial-temporal modeling of ENS data. Environmental ENS deployments will be used to illustrate the problems and methodology.

Relationship Between Prophets and Chapters in the Holy Quran

✱Walid K. Sharabati, George Mason University, Department of Mathematics and Statistics, Washington, DC 20016, *wsharaba@gmu.edu*; Yasmin H. Said, George Mason University; Edward J. Wegman, George Mason University

Key Words: Quran Social Network, Bipartite Network, 1-mode Network, Prophet Chapter Relationship, Prophets in Quran

In this paper, we analyze the social network of chapters in the Quran that mention prophets and their association. We further discuss how prophets and chapters were grouped in the Quran and identify the chapters that narrated stories about prophets. We used techniques from the network theory such as graphs and matrices and matrix algebra to model relations among chapters and prophets on three levels. On the first level, we investigate how prophets and chapters are related. Secondly, we investigate how chapters and chapters are related, and, finally, how prophets and prophets are related.

Distributions of Patterns and Statistics in Dynamic Conditional Random Fields

✱Donald E.K. Martin, North Carolina State University, 209-E Patterson Hall, 2501 Founders Drive,, Statistics Department, Campus Box 8203 NCSU, Raleigh, NC 27695-8203, *martin@stat.ncsu.edu*; John A.D. Aston, Warwick University

Key Words: Auxiliary Markov chain, classification, conditional random fields, discriminative models, undirected graphical models, hidden state sequences

This paper gives a dynamic programming method for computing distributions associated with a vector state sequence of a dynamic conditional random field. Distributions may be computed for such statistics as the number of segments of at least a certain length, the length of the longest segment, the number of occurrences of the various state categories, and also waiting time distributions of collections of patterns that must each occur pattern-specific numbers of times. Thus the methodology supplies extreme flexibility in data mining, while also integrating data mining into information extraction processes in a manner that correctly accounts for uncertainty in the hidden states, conditional on the observations. We show the the methodology may be applied to other graphical models as well.

318 General Issues in Highway Safety and Risk Assessment

Section on Risk Analysis

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

An Investigation of the Potential Effects of Nonresponse on the BTS Omnibus Household Survey

✱Pheny Z. Weidman, Bureau of Transportation Statistics, 1200 New Jersey Ave., S.E., Washinton, D.C., DC 20590, *pheny.weidman@dot.gov*

Key Words: Survey nonresponse, nonresponse bias, effect of survey nonresponse

The BTS Omnibus Household Survey (OHS) assesses the general public's opinion of the nation's transportation system by interviewing persons in randomly selected telephone households. The response rate for the OHS has been around 50% since its inception in 2000, raising concerns about the generalizability of its results. This study attempts to use more than one approach, e.g. comparing response rates on subgroups, analyzing estimates by level of effort, and etc., to investigate the potential effects of nonresponse on major OHS estimates. It also provides a short but comprehensive review of household survey nonresponse bias investigation methods. The results and methods used in this study can provide useful information to other government agencies regarding the conduct of nonresponse bias studies to meet the OMB's requirements for their own RDD surveys.

Electronic Records of Undesirable Driving Events

✱Oren Musicant, Ben-Gurion University of the Negev, CochavYair, Israel, *misucant@bgu.ac.il*; Hillel Bar-Gera, Ben-Gurion University of the Negev ; Edna Schechtman, Ben-Gurion University of the Negev

Key Words: Driver Safety, Driver Behavior, IVDR, Transportation

The cause for the majority of road crashes can be attributed to drivers' behavior. Recent in-vehicle monitoring technologies enable continuous and high resolution measurements of drivers' behavior. We analyze the information received from a novel in-vehicle technology which identifies the occurrence of undesirable driving events such as extreme braking and accelerating, sharp cornering and sudden lane changing. We use maximum likelihood estimation to fit a Negative-Binomial model to Events Frequency (EF) in a trip (N=100,013). Negative-Binomial regression results show higher EF in trip edges (trip beginning and trip end) than in the middle of the trip, higher EF for males than for females and higher EF at nighttime than at daytime. In addition, it shows a significant interaction between gender and time of the day-the increased EF at night was more extreme for males than females.

Recursive Partition for Early Outbreak Detection

✱Ross S. Sparks, CSIRO Mathematical and Information Sciences, Locked Bag 17, North Ryde, Sydney, International 1670 Australia, *Ross.Sparks@csiro.au*; Chris Okugami, CSIRO Mathematical and Information Sciences

Key Words: Monitoring, risk management, decision trees

Vehicle crash data in New South Wales (Australia) are collected using police reports for crashes with people injured or more than $500 worth in damages. Traditionally monitoring daily counts is used to identify unusual high counts and thus flag problems on the roads. However it is known that crashes start in small clusters. Technology for detecting spatial disease clusters can be found in Raubertas (1989) and Kulldorff (2001). A surveillance system which exploits the clustering nature of crashes by tracking the sources of variation will be demonstrated as an efficient way of monitoring crashes. This strategy is demonstrated to lead to earlier detection than monitoring counts aggregations over all dimensions. This strategy also has the advantage of describing how epidemics move over time in the population.

Evaluating the Effects of Commercial Vehicle Information Systems at Preventing Large Truck Crashes

✱Michael Anderson, Battelle, 2101 Wilson Blvd, Suite 800, Arlington, VA 22201, *andersonms@battelle.org*

Key Words: probability model, large truck crashes, roadside simulation, commercial vehicle information systems, relative risk

Commercial Vehicle Information Systems and Networks (CVISN) is a collection of information systems, communications networks, and Intelligent Transportation Systems that support commercial vehicle operations. CVISN includes information systems owned and operated by state agencies and the Federal government, private motor carriers, and other stakeholders. Two focuses of CVISN are: 1) The storing and sharing of safety information; and 2)Roadside electronic screening. A statistical analysis was conducted to measure the effects of these technologies on the safety of trucks and the general traveling public, through improved roadside enforcement. A probability model was developed to predict the number of large truck crashes, injuries, and fatalities that could be avoided under different simulated roadside enforcement scenarios that make use of various CVISN technologies.

Endogeneity of an Ordered Explanatory Variable in a Bivariate Model: Accidents and Deductible Choice for Young Drivers

Michel Grun-Rehomme, Université Paris2, ERMES; ✱Olga Vasechko, Research Institute of Statistics, 3 shota Rustaveli str., Kyiv, 01023 Ukraine, *O.Vasechko@ukrstat.gov.ua*; Noureddine Benlagha, Université Paris2, ERMES

Key Words: Bivariate ordinal model, automobile insurance, adverse selection, accident probability

Automobile insurers, using different observable explanatory factors to estimate the likelihood of accidents, have sought to better align premiums with accident probabilities. This raises the question of whether the insured drivers have changed their risk behavior. We analyze the coverage in the French automobile insurance market for information about the risk-aversion of drivers. This study looks at conditions of the insurance market. We consider four types of deductible and four frequencies of accidents (0, 1, 2 and more than 2). Autonomous estimates of the number of accidents may include an endogeneity bias. Introducing an accident variable with four modalities into the bivariate modeling enables us to test the significance of engine power in the equation of the coverage choice.

Addressing the Complexity of Speed Measurements at Intersections in Determining the Effectiveness of Speed-Reducing Treatments

✱Courtney Bokenkroger, Midwest Research Institute, 425 Volker Blvd, Kansas City, MO 64110, *cbokenkroger@mriresearch.org*; Karin M. Bauer, Midwest Research Institute

Key Words: transportation, speed data, ANOVA, repeated measures, before-after evaluation, distributional assumptions

The complexity of a statistical evaluation of the effectiveness of speed reduction treatments at high-speed intersections begins with first being able to understand and determine how transportation safety engineers typically measure effectiveness and extends into the complexity of the data collected and the distributional assumptions. In a recent study, vehicle speeds were measured at multiple locations along an intersection approach both before and after installation of the treatment. The analysis of the effectiveness of the treatment at a site was conducted using ANOVA with repeated measures. In this article, we discuss the complexity of the data and the factors that could and could not be accounted for in the analysis as well as their possible effects on estimates of the treatment effectiveness.

319 Statistical Applications in Engineering

Section on Physical and Engineering Sciences

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

A Cost-Efficient Approach to Wireless Sensor Network Design

✱Natallia V. Katenka, The University of Michigan, 439 West Hall, 1085 South University, Ann Arbor, MI 48109, *nkatenka@umich.edu*; Elizaveta Levina, The University of Michigan; George Michailidis, The University of Michigan

Key Words: wireless sensor network design, network cost, connectivity, coverage, random deployment, unreliable sensors

This study presents a general flexible approach for the wireless sensor network design under the random deployment mechanism. The cost of sensing and communications is incorporated into the design of the network, while allowing for unreliable sensors. The cost is treated generically and can correspond to either a fixed acquisition cost, or an operational cost or a combination of both. The main objective is to minimize the overall network cost, while enforcing the coverage and connectivity constraints. The proposed approach should be regarded as part of any feasibility study during the planning stages for the network deployment. An additional technical contribution is the derivation of a new simple bound on the probability of a network being connected, which exhibits a very good performance in simulations and unlike existing ones is shown to be better suited for network design studies.

The Conflict of Wave Function Collapse with Special Relativity: A Classical Statistician's View

✱John W. Sawyer, Simbalien, 2170 Ector Court, Atlanta, GA 30345, *zarnd@yahoo.com*

Key Words: wave function collapse, special relativity, EPR, Bernoulli variable

It is not widely appreciated that the way quantum mechanics actually produces data, wave function collapse, and special relativity do not sit well together. One problem arises because there is no absolute first measurement under special relativity when two measurements on coupled particles have spacelike separation. For the EPR thought experiment, spins are measured on an electron- positron pair. In this case the result of the conflict seems as if one tossed a fair coin with the left hand and simultaneously a biased one with the right, and the results always agreed on heads or tails. We might resolve this paradox by modeling observer knowledge rather than true physical outcome. Yet the question still remains: where/when in spacetime is the value of the Bernoulli random variable "electron(or positron) spin" actually realized?

Service-Level Agreement for the Stochastic Business Processes

✱Genady Y. Grabarnik, IBM T. J. Watson Research Center, 19 Skyline Dr., P.O. B. 704, Hawthorne, NY 10532, *genady@us.ibm.com*; Laura Shwartz, IBM T. J. Watson Research Center

Key Words: Stochastic business processes, SLA, Service Management

Stochastic business processes arise in the IT Service management. We consider service level agreement as a function on some of the statistics of the processes. SLA parameters characterizing shape of the function are defined to restrict losses of the service provider. We establish optimal boundaries for the parameters based on the condition that service provider is indifferent to the cost of SLA or the losses due to SLA breaches.

A Detailed Correlation Analysis of Enzymatic Reaction of a Single Protein Based on the Stochastic Network Model

✱Chao Du, Harvard University, Statistic Department, Science Center, 1 Oxford Street, Cambridge, MA 02138, *chaodu@fas.harvard.edu*; Samuel Kou, Harvard University

Key Words: Stochastic network model, autocorrelation function, single molecule, enzyme

The new advance in nanotechnology enables scientists to observe biological processes on a microscopic molecule-by-molecule basis. Recent single-molecule biophysical experiment on enzyme molecules shows that enzymes behave fundamentally differently from what the classical (Michaelis-Menten) model predicts. In the work by Kou, Xie et. al., a stochastic network model is proposed to explain the puzzle. Their detailed analysis, including the analysis of the first-passage-time distribution, shows that the network model is capable of explaining the experiment results. In this paper, further theoretical and data analyses are carried out to investigate/explain (i) the autocorrelation structure of the enzymatic reaction times, and (ii) the role of (substrate) concentration in enzymatic reactions. Our work shows that the stochastic network model is capable in explain the experimental data in depth.

Statistical Design and Analysis of Corrosion Experiments

✱Joanne R. Wendelberger, Los Alamos National Laboratory, Statistical Sciences Group, MS F600, Los Alamos, NM 87545, *joanne@lanl.gov*

Key Words: design of experiments, aging

Aging experiments are being conducted to investigate the initiation and progression of corrosion on metal plates. Experiment plans are developed to study the effects of variables thought to be potential contributors to corrosion. Statistical design techniques are used to develop plans that meet design objectives within specified resource limits on the number of runs. Data collected from the experiments will include both quantitative measurements and images. Analysis approaches are considered that will quantify the aging behavior and take into consideration the spatial heterogeneity of the corrosion behavior.

Statistical Adjustments to Engineering Models

✱Roshan J. Vengazhiyil, Georgia Institute of Technology, Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, *roshan@gatech.edu*; Shreyes N. Melkote, Georgia Institute of Technology

Key Words: Bayesian methods, Calibration, Computer experiments, Prediction, Validation

A common problem in engineering is that the models developed based on the underlying physics of the process do not always match satisfactorily with reality. On the other hand, pure statistical models fitted based on experimental data can give more realistic predictions, but can perform poorly when predictions are made away from the observed data points. This work proposes engineering-statistical models that combine the advantages of engineering models and statistical models. The engineering-statistical model is obtained through some adjustments to the engineering model using experimental data. The adjustments are done in a sequential way and are based on empirical Bayes methods. The usefulness of the methodology is illustrated using a problem of predicting surface roughness in a micro-cutting process.

Disaster Recovery Planning: SLA to Infrastructure Tradeoff

✱Laura Shwartz, IBM T. J. Watson Research Center, 19 Skyline Dr., Hawthorne, NY 10532, *lshwart@us.ibm.com*; Genady Y. Grabarnik, IBM T. J. Watson Research Center

Key Words: enterprise analytics, SLA, disaster recovery

Disaster recovery planning provides a capability to absorb the impact of an event and to prevent business failure. In this paper we focus on Immediate and Intermediate Recoveries which typically use a shared third-party facility that has IT Infrastructure. The recovery option is often dictated by industry standards, audit and legal requirements. Service Level agreements also provide valuable input to disaster recovery planning. By sustaining service levels, companies aim to mitigate such non-quantifiable cost as loss of client confidence and potential lost of sales and market share. High cost of recovery services provided by third-party makes advocating and planning disaster recovery a painful exercise. Without considering probability of the disaster, we are looking into optimizing the cost of a disaster recovery based on established guarantees and existing resources configuration.

320 General Methodology—Nonparametric and Semiparametric Style! ●▲

Section on Nonparametric Statistics, IMS
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Robust Likelihood Methods Based on the Skew-T and Related Distributions

Adelchi Azzalini, University of Padova; ✱Marc G. Genton, University of Geneva, Bd du Pont-d'Arve 40, Geneva 4, International CH-1211 Switzerland, *Marc.Genton@metri.unige.ch*

Key Words: Kurtosis, Multivariate, Robustness, Skewness

The robustness problem is tackled by adopting a parametric class of distributions flexible enough to match the behavior of the observed data. In a variety of practical cases, one reasonable option is to consider distributions which include parameters to regulate their skewness and kurtosis. As a specific representative of this approach, the skew-t distribution is explored in more detail, and reasons are given to adopt this option as a sensible general-purpose compromise between robustness and simplicity, both of treatment and of interpretation of the outcome. Some theoretical arguments, outcomes of a few simulation experiments and various wide-ranging examples with real data are provided in support of the claim.

Tests of Fit for the Beta Binomial Distribution

✱John Best, University of Newcastle, Australia, 12 Tobruk Ave, Carlingford, 2118 Australia, *Donald.J.Best@newcastle.edu.au*; John Rayner, University of Newcastle, Australia; Olivier Thas, Ghent University

Key Words: Goodness of fit, Moment tests, Empirical distribution function, Discrete distributions, Power studies

The beta-binomial has recently been suggested as an alternative model for analysis of simple market research preference tests (Bi and Ennis, 2001). Before accepting its use for this application and others it is sensible to test the fit of the model. Traditionally this can be done with a Pearson chi-squared test. Simonoff (2003) suggested a dispersion test while Lockhart et al (2007) discuss general tests of fit for discrete distributions based on the empirical distribution function. Skewness and kurtosis tests of fit will also be examined here. Preliminary results show the dispersion test has poor power for some alternatives, the Pearson test has some power for most alternatives, either the skewness or kurtosis test often has good power while the Anderson-Darling test is generally more powerful than the other tests for the alternatives considered.

Application of Deconvolution Method in a Graded Glucose Infusion Study

✱Bing Gao, Merck & Co., Inc, 126 E. Lincoln Ave, P.O. Box 2000, Rahway, NJ 07065, *bing_gao@merck.com*

Key Words: Deconvolution, Insulin Secretion Rate, Graded Glucose Infusion, C-peptide, Bell-Cell Sensitivity

Currently there are two ways to assess beta-cell glucose-sensitivity with regard to insulin secretion: the commonly accepted hyperglycemic clamp technique and the simpler oral glucose or meal challenge test. Yet the first is expensive with high requirements of clinical centers, and the second

has substantial variability. A graded glucose infusion (GGI) methodology comes as a compromise. Since beta-cells secret insulin and C-peptide with ratio 1, insulin secretion rate (ISR) can be estimated by deconvoluting C-peptide concentration using a two-compartmental model. To stabilize and avoid negative estimates of ISR, a regularized deconvolution method can be employed. The talk illustrates applying it in a real-world GGI study. The linear relationship was established between ISR and glucose concentration at various infusion rates, and the slope was used to assess beta-cell glucose sensitivity.

Empirical Investigation of Uncertainty Measures for Unbalanced Repeated Measures Models

✱Christine Spinka, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211, *spinkac@missouri.edu*; Scott Holan, University of Missouri

Key Words: Repeated Measures, Empirical Study, Unbalanced Design

Unbalanced repeated measures data can arise in many different areas of science, and many software packages now contain functions for model development and hypothesis testing for this type of data. Unfortunately, the methodology being used is often based on large samples and correctly specified models. These assumptions may be violated, creating the potential for a practitioner to draw questionable conclusions. Therefore, we propose a joint model of the mean and covariance for independently, but not identically distributed observation vectors. We consider a nonparametric mean function, and develop both bootstrap and asymptotic measures of uncertainty for unbalanced repeated measures. To illustrate our method, we present the results of a simulation study and data analysis that compare the bootstrap and asymptotic approaches in finite samples.

Assessing Individual Observer Agreement in Studies Involving Replicated Observations

✱Jingjing Gao, Emory University, Dept. of Biostatistics 3rd Flr, 1518 Clifton Rd NE, 318 RSPH, Atlanta, GA 30322-4201, *jgao@sph.emory.edu*; Michael Haber, Emory University

Key Words: coefficient of individual agreement, categorical data, repeated measurements

Methods for assessment of individual observer agreement are based on comparing disagreement between two observers to the disagreement between replicated observations made by the same observer on the same subject. These methods have been developed for continuous and binary observations. In this presentation we propose a new nonparametric approach to evaluating individual agreement between two observers making replicated categorical observations on each subject. We consider two situations: (1) a symmetric assessment of agreement between two observers, and (2) an assessment of the agreement of a new observer with an imperfect 'gold standard'. The reliability of the estimation method is examined via a simulation study. Data from a study aimed at determining the validity of diagnosis of breast cancer based on mammograms is used to illustrate the new concepts and methods.

On the Superposition of Overlapping Realizations of a Nonhomogeneous Poisson Process

✱Gustavo L. Gilardoni, Universidade de Brasilia, Departamento de Estatistica, Brasilia, DF, International 70910-900 Brazil, *gilardon@gmail.com*; Enrico A. Colosimo, Universidade Federal de Minas Gerais

Key Words: greatest convex minorant, nonparametric intensity estimation, optimal maintenance time, recurrent events, repairable systems

Many nonparametric procedures are well established when only one realization of a nonhomogeneous Poisson processes (NHPP) is observed. In this note it is argued that the case when several realizations are observed over different time intervals can be reduced to the former one. More precisely, it is shown that, when time is measured in the Total Time on Test scale, the superposition of overlapping realizations of a NHPP is also a NHPP and is sufficient for inferential purposes. Thus, using results already available when only one realization is observed, one can obtain for the many realizations setting, for instance, (i) the maximum likelihood estimator of a monotonic intensity or (ii) bootstrap confidence bands based on Kernel estimates of the intensity. The techniques are illustrated using a real data set containing the failure history of 40 power transformers.

Comparing the Efficiency of Automatic Forward Search Estimators with That of Alternative Robust Estimators in Linear Regression Models

✱Francesco Polverini, University of Florence, v.le Morgagni, 59, Dept. of Statistics, Florence, International 50134 Italy, *polverini@ds.unifi.it*; Bruno Bertaccini, University of Florence

Key Words: linear regression, robust estimators, automatic forward search, stopping rule, efficiency

In a previous paper we proposed a robust estimation of regression models using the so-called Automatic Forward Search (AFS). The proposed automatic, data-dependent stopping rule criterion, based on the elasticity of the estimates of the variance of the error term, s2, seems to work well. Our results showed clearly that the procedure produces estimators that are robust and far superior to the LMS estimators. We intend now to expand the research in at least two directions: AFS estimators will be compared with other robust estimators and some improvement in the stopping rule will be considered. Particularly when the data are not contaminated or the strength of the contamination is low, a stopping rule that takes into account the values of more than one statistic would probably eliminate the too early cut-off point that we found in some unusual samples.

321 Recent Advances in Regression and Linear Modeling

IMS, Section on Nonparametric Statistics
Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Recent History Functional Linear Models

✱Kion Kim, The Pennsylvania State University, 331B,thomas bldg, Pennsylvania State University, University park, PA 16802, *kion@psu.edu*

Key Words: functional linear model, historical linear model, varying coefficient model, basis expansion, B-spline basis

We propose a variant of historical functional linear models for cases where the current response is affected by the predictor process in a window into the past. Different from the rectangular support of functional linear models, the triangular support of the historical functional linear models and the pointwise support of the varying coefficient models, the current model has a sliding window support into the past. This idea leads to models that bridge the gap between varying coefficient models and functional linear (historic) models. The proposed estimation algorithm is shown to be fast in comparison to estimation procedures proposed for historical functional linear models, involving one dimensional basis expansions and one-dimensional smoothing procedures.

Estimating Expectations in Semiparametric Regression with Missing Responses

✱Ursula Müller-Harknett, Texas A&M University, Department of Statistics, 3143 TAMU, College Station, TX 77843, *uschi@stat.tamu.edu*

Key Words: empirical estimator, empirical likelihood, influence function, plug-in, semiparametric regression

We consider regression models with responses that are allowed to be missing at random. The models are semiparametric in the following sense: we assume a parametric (linear or nonlinear) model for the regression function but no parametric form for the distributions of the variables; we only assume that the errors have mean zero and are independent of the covariates. For estimating general expectations of functions of covariate and response (involving the mean response as a simple special cases), we use an easy to implement weighted imputation estimator. The estimator uses all model constraints and is therefore efficient in the sense of Hájek and Le Cam. The results are illustrated with examples.

Central Limit Theorems for Weighted Linear Processes and Their Applications

✱Kairat T. Mynbaev, Kazakh-British Technical University, Room 409, Tolebi 59, Almaty, 050000 Kazakhstan, *kairat_mynbayev@yahoo.com*

Key Words: central limit theorem, linear process, slowly varying regressor, time-series analysis, spatial model

We consider weighted sums of linear processes with weights arising in linear regressions with deterministic regressors. Central limit theorems are stated and applications are given to: 1) static models with slowly varying regressors and 2) two types of dynamic models (times series and spatial models). The emphasis is on the techniques of modeling deterministic regressors using functions of a continuous argument. The benefits of the approach are shown by comparing it with the recent paper by P.C.B. Phillips on regressions with slowly varying regressors. Only stable dynamic models are considered but the approach can be extended to nonstable ones.

Application of Directional Regression of EEG Data

✱Min Kyung Kim, The Pennsylvania State University, *mkkim1123@gmail.com*

Key Words: Dimension Reduction, Directional Regression, Principal Component Analysis

EEG is test that measures the electrical activity of human brains. An EEG data set that has motivated our research is concerned with the relation between genetic predisposition and alcoholism. There are two groups of subjects, an alcoholic group of 77 subjects and a control group of 45 subjects. Each subject was exposed to some pictures and the voltage values were measured from 64 sensors placed on the head through 256 time points. We apply Directional Regression (Li and Wang, 2007) to this data to find a dimension reduction space in which two groups are clearly separated by their means and variations. A difficulty is that the dimension of predictor is quite large, which makes it impossible to invert the sample covariance matrix. We combine Directional Regression with Principal Component Analysis as an intermediate reduction step to solve this problem.

Regression Level Set Estimation via Tree Warping

✱Gabe Chandler, Connecticut College, 18 Old Mill Rd, Quaker Hill, CT 06375, *gabriel.chandler@conncoll.edu*

Key Words: level, set, estimation, regression, dyadic, tree

Level set estimation for regression functions (i.e., the set on which the function exceeds a certain value) is an important problem with many applications. Willett and Nowak (2007) derived a estimation procedure that attains mini-max optimal rates of convergence among the class of dyadic trees. In practice, this method requires the selection of a smoothing parameter which determines the complexity of the tree. It is unlikely that the true level set will be well approximated by a union of large boxes. We propose a method that lessens the importance of the selection of this parameter by mapping a portion of the support likely to contain the boundary in such a way that the boxes induced by the dyadic tree better align with the boundary, thus allowing good approximation without the need for highly complex trees.

322 Analysis of Interval-Censored Data

Biometrics Section, WNAR

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Nonparametric Comparisons of Waiting Time Distributions in a Markov System Under Current Status Data

✱Ling Lan, University of Louisville, 101 Nob Hill Ln. Apt. 9, Louisville, KY 40206, *l0lan001@louisville.edu*; Somnath Datta, University of Louisville

Key Words: nonparametric test, multistate model, waiting time distributions, current status data

In this talk we present ways to construct test statistics comparing the weighting times in a given state in two or more multistate systems each of which has the same topology and follows a Markov transition times. The actual transition times are subjected to censoring resulting from a single random inspection time for each individual. Extensive power studies for our tests will be presented. An application to pubertal development data will also be considered.

Efficient Estimation for Bivariate Current Status Data with Proportional Odds Model

✱Bin Zhang, University of Missouri-Columbia, 510 High Street Apt. 316, Columbia, MO 65201, *bz38d@mizzou.edu*; Xingwei Tong, Beijing Normal University; Jianguo (Tony) Sun, University of Missouri

Key Words: Bivariate current status data, Copula model, Counting processes, Efficient estimation, Proportional odds model

Efficient estimation of the bivariate survival function from current status data is considered. The joint survival function is given by a copula model with the marginal proportional odds model. We are concerned with the maximum likelihood estimation (MLE) of the regression parameters and the association parameter based on the efficient score function, where the association parameter is dependent of the covariates. A large sample property shows that the estimators are consistent and asymptotically normal. Simulation studies show that the estimators have a good performance in both cases that the association parameter is dependent or independent of the covariates. A real data application is also studied to evaluate the proposed method.

A Multiple Imputation Approach to the Analysis of Current Status Data with the Additive Hazards Model

✱Ling Chen, University of Missouri-Columbia, Department of Statistics, 146 Middlebush Hall, Columbia, MO 65211, *lcbm9@mizzou.edu*; Jianguo (Tony) Sun, University of Missouri

Key Words: additive hazards model, current status data, multiple imputation, tumorigenicity experiment

This paper discusses regression analysis of current status data, which occur in many fields including cross-sectional studies, demographical investigations and tumorigenicity experiments (Keiding, 1991; Sun, 2006). For the problem, we focus on the situation where the survival time of interest can be described by the additive hazards model and a multiple imputation approach is presented for inference. A major advantage of the approach is its simplicity and it can be easily implemented by using the existing software packages for right-censored failure time data. Extensive simulation studies are conducted and indicate that the approach performs well for practical situations and is comparable to the existing methods. The methodology is applied to a set of current status data arising from a tumorigenicity experiment.

Inference and Uniqueness for Chain-of-Events Data Subject to Interval Censoring

✱Beth Ann Griffin, RAND Corporation, 1200 South Hayes Street, Arlington, VA 22202, *bethg@rand.org*; Stephen Lagakos, Harvard University

Key Words: chain of events data, uniqueness, survival analysis, estimation of distribution functions

Chain-of-events data result from periodic observations of a progressive disease process whose states occur in a prescribed order and are nonrecurring, with the last state being absorbing. Using a discrete time semi-Markov model, we develop an algorithm for nonparametric estimation of the distribution functions of sojourn times in a J state progressive disease model when transition times between states are independent and interval censored. Since issues of uniqueness for chain-of-events data are not well-understood, we also develop a series of sufficient conditions for uniqueness of the nonparametric maximum likelihood estimator, including situations where some but not all of its components are unique. The proposed methods are illustrated on a number of examples.

A New Approach to Test for the Homogeneity of Failure Time Distributions When Data Are Interval-Censored

✱HeeJeong Lim, California State University, Fullerton, McCarthy Hall #182J, 800 N. State College Blvd., Fullerton, CA 92834, *heejeonglim@fullerton.edu*

Key Words: Interval Censoring, Nonparametric Estimation

The log-rank test is most commonly used for right-censored data. It is known to be optimal under proportional hazards model, but fails to detect differences when the proportional hazards assumption is violated. To overcome this limitation, Lin and Wang (2004) proposed a new test procedure to compare two survival functions for right-censored data. Motivated by their idea, this article proposes a procedure to compare several distributions and extends it to interval-censored data. To obtain the test statistic for interval-censored data, the maximum likelihood estimate of the probability distribution is estimated using the EM algorithm. The simulation studies are conducted to evaluate the performance of the proposed the proposed test. This method is illustrated by analyzing a data set and the results are compared with other analyses that have been appeared in the literature.

Modeling Multivariate Current-Status Survival Outcomes with Frailty Model

✱Ying-Fang Wang, University of California, Davis, 5000 Orchard Park Circle, Apt. 5724, Davis, CA 95616, *yfwang@wald.ucdavis.edu*

Key Words: Frailty Model, Monte Carlo, EM algorithm

We first develop parametric modeling on clustered multivariate current-status survival data with frailty. Both inter-subjects correlations at the same state and intra-correlations between states are modeled by using multivariate exponential frailty distribution. Monte Carlo version of the EM algorithm and Newton-Raphson algorithm are used for parameter estimation. The model and computing method are then applied to analyze breeding success of the common raven.

323 Statistical Methods for Analysis of Microarray Data

Biometrics Section, WNAR

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Modeling Dependent Gene Expression

✱Donatello Telesca, The University of Texas M.D. Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 0447, Houston, TX 77030, *donatello.telesca@gmail.com*; Peter Mueller, M.D. Anderson Cancer Center; Giovanni Parmigiani, Johns Hopkins University

Key Words: Gene Expression, Markov Equivalence, POE Model, RJ-MCMC

We consider statistical inference for high-throughput gene expression data. Most traditional statistical methods implicitly assume independent sampling (conditional on some hyperparameters). Recognizing the limitations of independent modeling we develop a model that includes a simple dependence structure across genes. The important features of the proposed model are the ease of representing prior information on the nature of dependencies, model-based parsimonious representation of the signal as a ordinal outcome, and the use of a coherent probability model over both, structure and strength of the conjectured dependencies. Inference is implemented through a straightforward Markov chain Monte Carlo simulation, including posterior simulation over conditional dependence and independence.

The Beta-Binomial Distribution for Estimating the Number of False Rejections in Microarray Gene Expression Studies

✱Daniel L. Hunt, St. Jude Children's Research Hospital, Dept. of Biostatistics, 332 N. Lauderdale St., Memphis, TN 38105, *daniel.hunt@stjude.org*; Cheng Cheng, St. Jude Children's Research Hospital; Stanley B. Pounds, St. Jude Children's Research Hospital

Key Words: Beta-binomial, False discovery rate, Gene expression, Permutation

In analyses of microarray data, independence among gene expression levels is commonly assumed. This assumption implies that the number of false rejections V is binomially distributed, leading to a rough estimator of the false discovery rate (FDR), the empirical false discovery rate (eFDR). We assume that the number of false rejections V follows a beta-binomial distribution and construct the beta-binomial version of the empirical false discovery rate (bbFDR) as an estimator of FDR. This assumption accounts for the influence on V by correlated non-differentially expressed genes. We use a permutation procedure to generate realizations of V. In a simulation study, under several conditions, the bbFDR outperformed the eFDR in estimating the FDR. We also applied our method to an actual study comparing gene expression levels of soft tissue sarcoma samples and normal tissue samples.

Statistical Test for Clonality Using Array-CGH Data

✱Irina Ostrovnaya, Memorial Sloan-Kettering Cancer Center, 307 E. 63rd Str., New York, NY 10021, *ostrovni@mskcc.org*; Colin B. Begg, Memorial Sloan-Kettering Cancer Center; Adam Olshen, Memorial Sloan-Kettering Cancer Center ; Venkatraman E. Seshan, Columbia University

Key Words: tumors, cancer, clonal, metastases, copy number, aCGH

Genetic studies of pairs of tumor samples from the same patient can be performed to determine whether the tumors share a clonal origin, and the ultimate diagnosis may lead to different treatment options and prognosis. Array CGH data contain genomewide information on the copy number changes in the tumor cells and can be used to measure the similarity between the two tumors, which is quantified using endpoints of allelic changes and patterns of gains and losses. It is expected that recurrent changes specific to the disease may often appear in independent tumors. On the other hand, clonal samples might have different patterns due to accumulation of mutations after the tumors diverge and grow independently. These phenomena are taken into account in the new statistical approaches for assessment of clonality that we present. Both hypothesis testing and classification frameworks are considered.

A Markov Chain Monte Carlo Bayesian Network Approach To Infer Gene Co-Regulation Patterns

✱Younhee Ko, University of Illinois at Urbana Champaign, 1004 W Main St. #201, Urbana, IL 61801, *younko@uiuc.edu*; Chengxiang Zhai, University of Illinois at Urbana Champaign; Sandra L. Rodriguez-Zas, University of Illinois at Urbana-Champaign

Key Words: Bayesian network, Markov Chain Monte Carlo Gaussian mixture model, gene expression, gene network

Bayesian networks are a powerful framework to infer gene pathways from gene expression microarray experiments. A Markov Chain Monte Carlo (MCMC) implementation of Bayesian mixture networks was compared to an Expectation-Maximization (EM) implementation. Both implementations were applied to predict a signaling pathway using honey bee brain microarray data, and a cell communication pathway using mouse embryo microarray data. The MCMC estimates confirmed known gene relationships and uncover new relationships reported in the literature. The posterior probability distributions over candidate network structures had several local maxima. The lack of a single network structure clearly supported by the data may be due to the numerous conditions considered and limited information within condition. The posterior density estimates offered insights into the dynamic nature of gene networks.

Array Data Analyzer: A Novel Tool for Analyzing High-Dimensional Array Data

✱Jie Cheng, GlaxoSmithKline, 1250 S Collegeville Road, Mail code: UP-4335, Collegeville, PA 19426, *jie.j.cheng@gsk.com*; Joel Greshock, GlaxoSmithKline; Xiwu Lin, GlaxoSmithKline R&D; Kwan Lee, GlaxoSmithKline

Key Words: gene expression analysis, DNA copy number analysis, classification, biomarker discovery, predictive modeling, supervised learning

We developed an efficient Java based software tool that specializes in building and validating sets of molecular markers from whole genome array data. Comparing to other popular tools, our tool has two major advantages. First, instead of selecting markers using an arbitrary P value or FDR cutoff, our method finds the optimal set of markers based on cross validated model performance. Second, unlike other marker ranking mechanisms, which are always fixed tradeoffs between signal strength and signal consistency, our method can adapt to each individual dataset by finding an optimal tradeoff guided by cross validated model performance. Our tool has proven to be particularly effective for identifying amplified markers that are associated with a subset of cancers in a responding population. This technique is being effectively applied to most preclinical oncology assets at GSK.

Detecting Cellular Differentiation from Gene Expression Measurements

✱Feng Hong, University of Illinois at Urbana-Champaign, 101 Illini Hall, MC-374, 725 South Wright Street, Champaign, IL 61820, *fenghong@uiuc.edu*; Xuming He, University of Illinois at Urbana-Champaign; Sheng Zhong, University of Illinois-Urbana Champaign

Key Words: microarray, cellular differentiation, mixture

Cellular differentiation refers to the biological process by which unspecialized cells becomes specialized into one of the many cells. As a result, cells that start from a homogeneous population end up at a mixture of different cell types. We investigate the possibility of detecting cells homogeneity based on gene expression measurements. Because the microarray gene expression measurements are usually obtainable only at the cell population level, it is not easy to know whether and how they can detect mixtures in a population. In this paper, we rely on time course data and propose a statistical approach to identification of cellular differentiation. We tested the proposed method on a synthetic dataset as a validation study. We also applied the method on a gene expression time course dataset from a mouse stem cell differentiation experiment. The results are encouraging in both studies.

324 Use of Decision Analysis in Clinical Trials ●▲

Biopharmaceutical Section, Section on Health Policy Statistics, Biometrics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Benefit-Risk Assessment in Dose-Finding Trials

✱Hui Quan, sanofi-aventis, BX2-403A, 200 Crossing Boulevard, PO Box 6890, Bridgewater, NJ 08807, *hui.quan@sanofi-aventis.com*; Yujun Wu, sanofi-aventis

Key Words: monotonic assumption, restricted estimate, multiplicity, confidence interval

The objective of dosing finding trials is to find not only the effective but also safe doses for either future late phase trials or the ultimate patients. Most methodologies developed for analyzing data from dosing finding trials mainly focus on finding the effective doses or minimum effective dose through hypothesis testing. In this research, we will focus on benefit-risk assessment mainly through point estimation. To gain efficiency, monotonic dose-response relationship at least for the safety parameter will be assumed. However, the monotonic assumption cannot be made directly for the benefit-risk measures. This will not permit the use of step down procedure for multiplicity adjustment. Several different approaches will be discussed and compared. Data examples will be used to demonstrate the properties of these approaches.

Decision Analysis: A Tool for Biostatisticians To Select and Present Clinical Trial Design

✱Marc L. de Somer, McMaster University, 77 Dromore Crescent, Hamilton, ON L8S 4B1 Canada, *desomer@mcmaster.ca*; Lehana Thabane, McMaster University

Key Words: clinical trial, adaptive design, fixed design, ATD, RCT, decision analysis

Unlike traditional fixed clinical trial design, adaptive trial design (ATD) enables learning and adaptation during ongoing trials without compromising their integrity. Substantive ethical, economic and scientific benefits can accrue to a wide range of stakeholders - from patients and investigators, to sponsor and investors, ethics boards and regulatory authorities. Many influential stakeholders are unfamiliar with ATD and are concerned about the increased complexity, time, risk and cost, as well as acceptance by regulators and non-expert stakeholders. Decision analysis can help biostatisticians make a rational and persuasive case for adopting ATD in clinical trials because the improved benefit-risk of ATD often outweighs the real or perceived cost, time or risk increment. We present a case study of decision analysis for alternative RCT designs.

A Novel Approach to Proof-of-Concept in Stroke Recovery

✱Christopher Assaid, Merck Research Laboratories, UG1C-46, P.O. Box 1000, North Wales, PA 19454, *Christopher_assaid@merck.com*

Key Words: Stroke Recovery, Adaptive Design, Sample Size Re-assessment

Recent failures in the identification of effective treatments for neurologic recovery from acute stroke, including late-stage confirmatory trials, have increased the urgency for progress in the development of pharmaceutical candidates in this therapeutic area. Stroke recovery trials are not only expensive, but they pose numerous challenges, from both a clinical and statistical standpoint. We propose a proof-of-concept model for the development of stroke-recovery (or neuroprotective) drug candidates that is novel on a number of different dimensions, including the primary efficacy endpoint as well as adaptive elements of trial design (including a futility assessment as well as a sample size re-assessment). In addition, we propose an aggressive utilization of a group-sequential method for monitoring of all-cause mortality, a key safety measure in trials of this nature.

A Bayesian Approach To Assess Probability of Success of a Study

✱Fang Liu, Merck Research Laboratories, 126 E. Lincoln Ave, PO BOX 2000, Rahway, NJ 07065, *fang_liu@merck.com*

Key Words: Bayesian, Power, Probability of success (POS)

Power is one way to assess the POS of a study. If the alternative hypothesis (H_1) of a study is composite, power is often provided for a range of plausible values of the true parameter under H_1 to account for our uncertainty about the parameter. When some prior knowledge I (from literature or pilot studies) exists about the parameter, it can be formalized in a Bayesian framework with a reasonable probability distribution of the parameter. Thus, an alternative quantification of POS is the expected value of the power function over the parameter space under H_1 given I. Monte Carlo simulations can be used if closed form integration does not exist. This approach would have the advantages of easy interpretation and convenient comparisons of POS, and ultimately, facilitating easier and quicker decisionmaking. A real-world example will be used to illustrate the new approach.

Using Posterior Probability as Decision Rule for Nonsuperiority Study

✱Danielle Sheng, Merck & Co., Inc, Merck & Co., RY34-A316, Rahway, NJ 07065, *danielle_sheng@merck.com*; Matilde Sanchez, Arena Pharmaceuticals

Key Words: Posterior Probability, Non-Superiority, Interim Analysis

In clinical trials, when previous studies showed conflicting efficacy results of the two doses (e.g., Dose A was superior to Dose B in one study but was inferior to Dose B in another study), it is of interest to confirm whether or not the higher dose is superior to the lower dose, in order to make a go/no-go decision on the development of the higher dose. Similar to inferiority testing, a non-superiority testing framework can be constructed. Interim analyses can be used to increase the efficiency of the study, and a posterior probability can be used as the decision rule for interim analysis because clinicians easily interpret it. This talk will discuss the framework of non-superiority design using posterior probability for interim monitoring and the operating characteristics of such a design will be demonstrated through simulations.

Hypothesis Testing with a Buffer Area

✱Boshao Zhang, Medical University of South Carolina, 135 Cannon Street, Suite 303, Charleston, SC 29403, *zboshao@yahoo.com*

Key Words: sample size, hypothesis testing, clinical significance, quality control

A testing strategy is proposed to avoid the ambiguity of a "negative result." It distinguishes important effects from those in a buffer area. If a result falls into the buffer area, the effect can be declared positive with a probability, or declared as pending. There may be clinical uses. For example, a new blood pressure lowering drug is tested with mean 145 mm Hg, sd 13 mm Hg for the study group. A buffer is set for reduction between 4 and 10 mm. Sample size can be calculated for conventional testing, but a test with a buffer area yields a markedly reduced sample size.

Decision Analysis in Late-Stage Clinical Trials

✱Suvajit Samanta, Merck Research Laboratories, 126 E. Lincoln Avenue, P.O. Box 2000, Rahway, NJ 07065, *suvajit_samanta@merck.com*; Weili He, Merck Research Laboratories

Key Words: clinical trials, decision rule, robust estimator, Bayesian

In late stage clinical development (mainly in phase III), an organization often has to select the optimal strategy among many for the development of a particular drug or identify the most prosperous candidates. In principle, different decision methods such as Expected Utility rule, Prospect Theory, Risk Value rule, and various other decision methods can be used for this purpose. We now extend the existing decision methods in both Bayesian and non-Bayesian setup to incorporate the issues in clinical trials into the decision model. We conduct extensive simulation studies to compare different decision methods including the newly proposed methods in various clinical trial settings. This is an ongoing project and our findings in the simulation studies will be included in the talk.

325 Contributed Poster Presentations

Biopharmaceutical Section, Biometrics Section, Section on Health Policy Statistics, Section on Physical and Engineering Sciences, Section on Risk Analysis, Section on Statistical Computing, Section on Statistical Consulting, Section on Statistics and the Environment, Section on Statistics in Epidemiology, Social Statistics Section

Tuesday, August 5, 2:00 p.m.–3:50 p.m.

Parameter Estimation in Astrophysical Accretion Disc Models

✱Elizabeth Martínez-Gómez, National Autonomous University of Mexico, Institute of Astronomy, Circuito de la Investigación Científica S/N, Del. Coyoacán, Mexico, International 04510 Mexico, *affabeca@gmail.com*; Jorge Cantó-Illa, National Autonomous University of Mexico; Victor M. Guerrero Guzmán, Instituto Technológico Autónomo De México

Key Words: parameter estimation, maximum likelihood, least-squares, genetic algorithm, Astronomy

Parameter estimation is a discipline that provides tools for the efficient use of data for aiding in mathematically modeling of phenomena and the estimation of constants appearing in these models. In astrophysics, as well as in other fields, there are many ways (mathematical and statistical) to estimate the parameters in a model, but most of them require the calculation of derivatives, and they are computational expensive. In other words, the parameter estimation can be treated as an optimization problem. The most commonly used techniques include least squares (linear, orthogonal, gradient-weighted), Kalman filtering, robust techniques (clustering, regression diagnostics), and maximum likelihood estimation. In this work, we estimate the parameters involved in a series of accretion disc models around pre-main sequence stars through the implementation of a "modified" genetic algorithm.

Estimating Critical Micelle Concentration (CMC) via Change-Point Regression Modeling

✱Tom Filloon, Procter & Gamble, 8700 Mason-Montgomery Road, Mason, OH 45036, *filloon.tg@pg.com*

Key Words: change-point, segmented regression

In many regression modeling applications, a simple polynomial or parametric form of model is adequate for fitting a particular data set. In surfactant experiments, piecewise (aka segmented) polynomial regression is needed to yield good estimation of a surfactant's critical micelle concentration (CMC), in other words change-point. Such model fitting is nontrivial but can be performed via standard nonlinear curve fitting software.

Building Bridges Between Planetary Sciences and Statistics

✱Padma Yanamandra-Fisher, Jet Propulsion Laboratory, California Institute of Technology, *amy.braverman@jpl.nasa.gov*; David Holt, University of California, Irvine; Ka Wong, University of California, Los Angeles; Amy Braverman, Jet Propulsion Laboratory; Sean Jezweski, California Institute of Technology

Key Words: PCA, Planetary Sciences, Statistical Outreach

We present an example of applying statistical methods to large volumes of multi-variate, multi-spectral remote sensing planetary data of solar system objects. An immediate goal for research is to identify various trends of the underlying latent variables of the physical system and relate them to the observed data. In particular, since 2005, Juptier has been undergoing a global upheaval, with well-known white oval changing color; outbreaks at various latitudes, and their subsequent development and decay over varying timescale highlight one important question for Jovian science: What causes global upheavals? Although a difficult question to answer, we can break it down to a simple inquiry: Are the various outbreaks similar in morphology and composition? We apply Principal Component Analysis (PCA) to these features and explore the results of the analyses of our pilot project to models.

Algorithm of Periodic Signal Reconstruction from Noisy Environment

✱Amy L. Potrzeba, University at Albany, School of Public Health, 1 University Place, Rensselaer, NY 12144, *ap412489@albany.edu*; Igor G. Zurbenko, University at Albany

Key Words: adaptive spectral estimation, filtering of time series, KZFT

Method of reconstruction of signals with concentrations in the different frequencies has been developed and implemented into software algorithms: R-package KZFT, http://cran.us.r-project.org. In practice some signals have definite frequency lines synchronized by specific physical periodicities. We develop method of filtering such specific frequencies from a noisy environment and reconstruction of the signal due to those frequencies. Estimation of accuracy of such reconstruction is provided. Intensive simulations methods are used to provide detection limits for such reconstruction and best choice of parameters to use. Adaptive spectral estimation from KZFT over hourly record of atmospheric pressure reviles distinctive spectral lines with periods easily recognizable as Moon and Sun influences. Those lines produced strictly periodic component which can be detected by KZFT.

A Propensity Score Analysis To Improve Covariate Balance in a Randomized Experiment: Reducing Bias in the Prevention of Perinatal Sepsis (PoPS) Trial

✱Cassandra K. Wolos, Harvard University, 1 Oxford Street, Statistics Department, Science Center, Cambridge, MA 02135, *wolos@fas.harvard.edu*; Donald B. Rubin, Harvard University

Key Words: Propensity Scores, Randomized Experiments, Balanced Distribution

This work applies propensity score methods, developed for observational studies, to a randomized experiment to create better covariate balance between treatment groups than expected by chance. The PoPS trial was designed by the Center for Disease Control and South African investigators to evaluate the effect of chlorhexidine wipes on the prevention of early onset neonatal sepsis. 8005 South African mother-newborn pairs were randomly assigned to treatment or placebo at the time of delivery. Covariates characterize medical history and delivery. Though randomization ensures covariate balance in expectation, chance imbalances were suspected, and overall rate of neonatal sepsis drifted over time. By estimating propensity scores and comparing treatment groups within propensity score subclasses, we address the time drift and reduce covariate imbalance importantly without involving outcome data.

Investigating the Use of Computational Algorithms for Constructing Nonregular Robust Parameter Designs

Debra Ingram, Arkansas State University; ✱Latia Carraway, Arkansas State University, Department of Mathematics and Statistics, P.O. Box 70, State University, AR 72467, *latia.carraway@csm.astate.edu*

Key Words: robust parameter design, non-regular design, fractional factorial design, minimum aberration

The general robust parameter design (RPD) problem involves finding the settings of the control factors that minimize process variation due to changes in the (uncontrollable) noise factors. In recent work, a minimum aberration criterion for ranking nonregular robust parameter designs was proposed and optimal non-regular RPDs were constructed by exhaustive search. The purpose of this work is to investigate the use and performance of efficient computational algorithms for the construction of optimal nonregular robust parameter designs.

Response Surfaces, Blocking, and Split Plots: An Industrial Experiment Case Study

✱Willis Jensen, W.L. Gore & Associates, 3750 West Kiltie Lane, Flagstaff, AZ 86003-2400, *wjensen@wlgore.com*

Key Words: Blocking Factor, Central Composite Design, DOE, Split-Plot Design

The particular experiment performed in this case study has response surface, blocking and split-plot elements which pose challenges when doing the analysis. The objective of the experiment was to optimize the factor levels in terms of two different responses to ensure that the manufactured product meets the specifications. We first illustrate and discuss the original, incorrect analysis that was done and then show an alternative analysis that takes into account all the elements. We make some comments on the importance of tailoring the analysis to match the way the design was done.

Risk and Protection Factors in Breast Cancer

✱Luis S. Cid, University of Concepción, Dept. of Statistics, Concepcion, International 01 Chile, *lucid@udec.cl*; Jorge P. Pierart, Universidad de Concepcion; Sandra Ramirez, Universidad de Concepcion

Key Words: Breast cancer, risk factor, survival

Breast cancer is among the main causes of death of women and there are several well known risk factors. In this study we explore the relationships between some of these factors and the life expectation after diagnostic. We studied a sample of 1200 women with breast cancer to determine the relationships between the some of the risk factors and the expected survival time, using Kaplan and Meier Product Limit Estimator, Cox's Proportional Hazard models and Median Survival Time, among other statistics, to estimate the expected survival time for different risk groups. Results show that while some risk factors provide shorter life expectation (age at first pregnancy, age of menopause), other factors act in the opposite way (age at detection), that is, as protection factors. We run separate studies for early stages (I and II) and advanced stages (III and IV) of the disease.

A Simulation Study of Performance of Hypertabastic and Hyperbolastic Survival Models in Comparison with Classic Survival Models

✱Zoran Bursac, University of Arkansas for Medical Sciences, 4301 W. Markham, Slot 781, Little Rock, AR 72205, *zbursac@uams.edu*; Mohammad Tabatabai, Cameron University; D. Keith Williams, University of Arkansas for Medical Sciences

Key Words: Hypertabastic model, Hyperbolastic model, parametric survival models

We conduct a simulation study to evaluate the performance of two recently introduced parametric survival models, which we call hypertabastic and hyperbolastic, respectively. These models are very flexible and can fit various hazard shapes. Here we sample from Weibull, log-normal and gamma distributions, while varying sample size, proportion of censored observations and parameter combinations of the specified distributions. Since all distributions under consideration have two parameters we use the average -log likelihood as an overall measure of goodness-of-fit, based on 1,000 simulations for each condition. We compare the hypertabastic and hyperbolastic survival models with commonly used Weibull, log-logistic and log-normal survival models. Results show that hypertabastic and hyperbolastic survival models are to certain extent, robust to variations in the underlying distribution.

Power and Sample Size Calculations for Current Status Survival Analysis

✱John M. Williamson, Centers for Disease Control and Prevention, 4770 Buford Highway, MS F-22, Atlanta, GA 30341, *jow5@cdc.gov*; Hung-Mo Lin, Mount Sinai School of Medicine; Hae-Young Kim, New England Research Institutes

Key Words: Current status data, Power, Sample size, Survival analysis

Although sample size calculations have become an important element in the design of research projects, such methods for studies involving current status data are scarce. Here we propose a method for calculating power and sample size for studies using current status data. This method is based on a parametric survival model that assumes proportional hazards and a two-group comparison that uses the Wald test. As expected, studies with current status data have substantially less power than studies with the usual right-censored failure time data. Our simulation results demonstrate the merits of these proposed power calculations.

Multiple-Event Survival Analysis of Chronic Stable Angina and Asymptomatic Known Coronary Artery Disease

✱Zugui Zhang, Christiana Care Health System, 131 Continental Drive, Suite 202, Newark, DE 19713, *ZZhang@ChristianaCare.org*; Paul Kolm, Christiana Care Health System; Ewen Edward, Christiana Care Health System

Key Words: multiple-event data, unconditional multiplicative hazards model, conditional multiplicative hazards model, chronic stable angina, asymptomatic known coronary Artery Disease, correlation of multiple events

In coronary studies, multiple failure time data frequently arise since patients having one cardiovascular disease (CVD) event likely had a greater chance of subsequent events. In this study, given that more than half of events were recurrent, multiple-event survival analyses were conducted to examine the risk of multiple CVD events in the study population. Results from three types of analyses are compared: the traditional survival analysis, unconditional multiplicative hazards model assuming independence of multiple events, and the conditional multiplicative hazards model assuming correlated multiple events. The unconditional analysis using the Anderson-Gill multiplicative model resulted in different trends of increasing risk with increasing blood pressure category, demonstrating the necessity of considering the correlation of multiple events.

A Cure Model for Alternating States Data with Frailty

✱E. Paul Wileyto, University of Pennsylvania, 3535 Market St., Suite 4100, Philadelphia, PA 19104-3309, *epw@mail.med.upenn.edu*; Daniel F. Heitjan, University of Pennsylvania

Key Words: multivariate survival, cure models, alternating states, within-subject correlation, frailty, repeated measure

Alternating states data arise in behavioral medicine contexts such as smoking cessation clinical trials. Current abstainers may remain abstinent or lapse into smoking, whereas current smokers may abandon the quit attempt or recover abstinence. Individuals may transition between smoking and abstinence several times in a single trial. One can describe the lengths of sojourns in these states with cure models, which include a latent variable to represent membership in the cured fraction (permanent abstinence or abandonment of the quit attempt) together with a survival model for the duration of the sojourn in the noncured fraction. We parameterize cure fractions by logistic regression, non-cured transition hazards by Weibull models, and within-subject correlation by a shared inverse-Gaussian frailty.

Weight Selection in Survival Analysis with Delayed Effect

✱Boguang Zhen, U.S. Food and Drug Administration, 1401 Rockville Pike, Rm 217S, HFM-219, Rockville, MD 20852, *boguang.zhen@fda.hhs.gov*

Key Words: Delayed Effect, Survival Analysis, Therapeutic Tumor Vaccine

Therapeutic tumor vaccines hold great potential for use in treating cancer. As this novel therapy moves from laboratory findings to clinical practice, we face new issues and uncertainties involving study design, efficacy, and long-term safety. One of these issues is called delayed effect. Unlike chemotherapy in which study drugs have direct impact on cancer, it usually takes several weeks or months for a therapeutic tumor vaccine to mount an effective immune response. The impact of delayed effect in survival analysis should be considered. This talk will explore analysis methods for analyzing time-to-event data when delayed effect exists, especially for choosing proper weights in the analysis. A real example will be presented, and simulations will be used in weight selections. Results from using different analysis methods will be compared and discussed.

Calculation of Sample Size for Clinical Trials in Presence of Dependent Right Censoring

✱Huang Ching-yu, National Health Research Institutes, Keyan Road, Zhunan Town, Miaoli county, 350 Taiwan, *cyhuang.st88g@nhri.org.tw*; Peng Nanfu, National Chiao Tung University; Chang Hsing-Yi, National Health Research Institutes

Key Words: conditional continuous time Markov chain, dependent right censoring, Gumbel-Barnett copula

In this thesis, we propose a new model, conditional continuous time Markov chain, modeling the process of clinical trials in presence of dependent right censoring. Based on this model, we can calculate the sample size to satisfy the given significance level and power. We use the log-rank test to detect the difference of treatment effects. This work extends the method proposed by Lakatos. We use the Gumbel-Barnett copula to model the dependent structure of survival time and censoring time. Using this new method, we find that as the value of this copula parameter increase the more sample size we need.

Comparing Survival Functions Between a Surgical Cohort and Matched Members of the General Population

✱David M. Thompson, University of Oklahoma, 825 NW 49th St., Oklahoma City, OK 73118, *dave-thompson@ouhsc.edu*; Ronald Elkins, University of Oklahoma

Key Words: matching, survival analysis

A study of survival following a surgical procedure motivated a comparison between patients in a surgical cohort and age- and sex-matched members of the overall U.S. population. Because surgeons performed the procedure for several years, we also matched each patient on the year of surgery by using the standard U.S. actuarial table for that year. The approach permitted us to aggregate, from patient-matched survival functions, the expected probability of survival beyond the year of surgery for age-, sex-, and year-matched persons in the general population. Following a decline associated with early hospital deaths, patients' survival appeared to parallel that of the general U.S. population to 13 years.

Software To Support Weibull Inference

✱John I. McCool, Pennsylvania Great Valley, 30 E. Swedesford Road, Malvern, PA 19355, *mpt@psu.edu*

Key Words: Weibull, Type II Censoring, Pivotal Functions, Confidence Intervals, Hypothesis Tests, Prediction Intervals

Confidence Limits and Hypothesis tests on the Weibull shape parameter and percentiles based on Type II censored samples require the distribution of pivotal functions which must be obtained by Monte Carlo simulation. A software program named Pivotal has been developed in Visual Basic for computing these distributions for the user's choice of sample size and censoring number. The program generalizes to include series systems of m identical components. This generalization can be exploited for analyzing sudden death tests. Another program option lets the user write the values of the estimated shape and scale parameters to a file. These output values can be imported into a spreadsheet or statistical package and modified to permit the computation of confidence limits on reliability, prediction intervals, and OC curves for hypothesis tests about a Weibull percentile or shape parameter.

Maximum Likelihood Estimation Methodology Comparison for the Three-Parameter Weibull Distribution with Applications to Offshore Oil Spills in the Gulf of Mexico

✱William V. Harper, Otterbein College, Mathematical Sciences, One Otterbein College, Westerville, OH 43081-2006, *wharper@otterbein.edu*; Thomas R. James, Otterbein College; Ted G. Eschenbach, TGE Consulting

Key Words: Weibull, Maximum Likelihood, reliability

Maximum likelihood estimation of the two parameter Weibull distribution is straightforward. However, there are multiple methods for maximum likelihood estimation of the shift parameter of a three parameter Weibull. This paper evaluates MLE methods using oil spill data from the Gulf of Mexico. Recommendations are made for practice, and Excel VBA routines have been developed that allow practitioners to implement the recommended practices. These routines are available free at http://faculty.otterbein.edu/Wharper.

Causal Modeling Approach to Differential Treatment Effects

✱Jennifer Faerber, University of Pennsylvania, 3700 Walnut Street, Philadelphia, PA, *jfaerber@dolphin.upenn.edu*; Thomas Ten Have, University of Pennsylvania; Dylan Small, University of Pennsylvania; Marshall Joffe, University of Pennsylvania

Key Words: structural mean models, causal modeling, IV estimation, g-estimation, sequential ignorability, depression

We examine differential treatment effects by employing a linear rank preserving model approach for analyzing the effect of a randomized baseline intervention modified by a post-randomization factor on a follow-up outcome. Unlike standard interaction analyses, our approach does not assume that the post-randomization factor is randomly assigned to individuals (a form of sequential ignorability). However, there is a tradeoff with other assumptions. Consistent estimation of causal effects without sequential ignorability employs weights under the G-estimation approach that are optimal in terms of semi-parametric efficiency but under sequential ignorability. In this context, we will present simulations and analyses of the effect a cognitive behavioral therapy intervention on depression, where the treatment effect is modified by behavioral factors evolving during the treatment.

Type I Error Rates and Power for Factor Analysis Applied to Data from the Latent Class Model

✱James P. Long, Columbia University, 5572 Lerner Hall, New York, NY 10027-8356, *jpl2116@columbia.edu*; John Kolassa, Rutgers, The State University of New Jersey

Key Words: factor analysis, maximum likelihood factor analysis, latent class model, latent class analysis, simulation

Researchers frequently use goodness of fit statistics from factor analysis applied to categorical data despite model assumptions requiring multivariate normal manifest variables and continuous latent variables. Violating model assumptions may result in erroneous type I error rates and suboptimal power when conducting hypothesis testing regarding the number of latent variables. We analyze type I error rates and power generated by applying likelihood ratio tests appropriate for factor analysis to data generated from a latent class model. Our results indicate that despite model misspecification, type I error remains close to nominal levels. Additionally power produced by factor analysis is roughly equivalent to power produced by latent class analysis.

Smaller but More Unequal World? A Network Analysis of the World City System

✱Xiulian Ma, The University of Utah, 779 Heritage Center, Salt Lake City, UT 84112, *xiulian.ma@hsc.utah.edu*

Key Words: Network Analysis, Longitudinal Network Modeling, World Cities, Global Cities, City System

Cities started to organize into a system on the world level. This paper will examine several networks of the world city system, in which cities are linked through air passenger flow, internet backbone, or transnational corporation headquarters and branch location. The study will employ network analysis (such as scale-free Modeling and longitudinal modeling) and generalized estimation equations. First, it will answer questions that so far have been ignored by world city researchers, such as -What kinds of network are the world city network? What are their growth tendency and mechanism? Then, the study will explore two questions pertaining globalization (1)do various world city networks have presented us a "smaller" (more connected) but more unequal world? (2) To what extent does inequality in the world system still determine the inequality in the city system?

Factorial Invariance and Robustness to Low Variability: Maximum Likelihood Factor Analysis vs. Correlation Constraint Analysis

✱Rochelle E. Tractenberg, Georgetown University, Dept of Neurology, Suite 207 Building D, 4000 Reservoir Rd., NW, Washington, DC 20057, *ret7@georgetown.edu*

Key Words: factor analysis, tetrad, causal modeling, factorial invariance, latent structure

This poster describes a planned simulation study comparing methods to uncover latent structure. Exploratory factor analysis (EFA) will be contrasted with the more explicitly causal correlation constraint analysis (CCA). The project will compare CCA and EFA in terms of accuracy, invariance, and sensitivity to variability within 100 simulated samples built to specifications and repeated with high and low levels of variability. The study will test whether: 1. CCA is more sensitive to lower levels of variability than EFA. 2. CCA results are more factorially-invariant than EFA results. The proposal will determine if CCA correctly recovers the 'true' latent variables and structure from simulated observed data more consistently than EFA ("accuracy"), recover the same 'true' model 95% of the time ("invariance"). Replication of results in samples with high/low variance will support sensitivity.

Classification and Regression Tree (CART) Analysis: Advantages on Interaction Effect

✱Yinmei Huang, The University of Akron, Department of Sociology, Akron, OH 44325, *huang_ym@hotmail.com*

Key Words: CART, Interaction Effect

The purpose of this paper is twofold. Firstly, introduce CART (classification and regression tree) with a global overview from a mathematical basis to its application, both conceptually and statistically. Next, respectively practice two kinds of tree analysis: classification, and regression tree, focusing on its advantages of disclosing and explaining interaction effect by comparison with conventional methods, such as OLS regression and logistic regression. Data in this investigation come from the National Survey of Families and Household 1988.

Chinese and American Self and Other: Calibration in Relative Rank on Everyday Tasks

✱Adam Molnar, Bellarmine University, 2001 Newburg Road, Louisville, KY 40205, *amolnar@bellarmine.edu*; Shali Wu, The University of Chicago

Key Words: intercultural communication, Chinese, psychology, percentile, standardization, regression

Cultures differ in the construal of self and other. Research suggests that Americans focus internally, while East Asians view self and other as intricately bound. If true, Chinese should have the perspective to calibrate relative societal rank more accurately than Americans. Students in both countries were asked to compare themselves to their peers on common tasks such as throwing a baseball and being on time. The psychology experiment had several interesting issues in data analysis. People tend to express percentile ranks in deciles, making the data nonregular and complicating standardization. After accounting for clumps and finding a standard distribution, regression suggested that information matters. When information is absolute and focal, both cultures calibrate well, but when information is relative and contextual, Chinese are more accurate than Americans.

Studying Mexican Return Migration at the Locality Level in 2000 and 2005

✱Claudia Masferrer, The University of Texas at Austin, 1 University Station, G1800, Population Research Center, Austin, TX 78712, *cmasferrer@math.utexas.edu*; Bryan Roberts, The University of Texas at Austin

Key Words: Return International Migration, Poverty, Mexico

Mexican migration to the United States has been a complex historical phenomenon with various implications (economical, social, political, demographical, etc.) on both sides of the border. In this context, the comprehension of return migration is of great importance. Using the Mexican 2000 and 2005 Population censuses, we study Mexico-U.S. return migration at the locality level for detecting those characteristics that make them have a higher propensity to receive migrants back, and relates return migration with poverty and exclusion to analyze how structural factors at the community level affect the intensity of return migration. The study uses Logistic Regression and Generalized Linear models to show that urban communities with lower levels of poverty and social exclusion are more likely to receive migrants back. It also notes some changes in 2000 and 2005.

Social Network Analysis (SNA) of Weighted Telecommunication Graphs

Norbert Walchhofer, Vienna University of Economics and Business Administration; ✱Angela Bohn, EC3 E-Commerce Competence Center, Donau-City-Str. 1, Vienna, 1220 Austria, *angela.bohn@gmail.com*

Key Words: Social Network Analysis (SNA), closeness centrality, telecommunications graphs, weighted graphs, Island algorithm, graph clustering algorithms

SNA provides a wide range of tools that allow examination of telecommunications graphs. Those graphs contain vertices representing cell phone users and lines standing for established connections. Many SNA tools do not incorporate the intensity of interaction. This may lead to wrong conclusions since the difference between best friends and random contacts can be defined by the accumulated duration of talks. To solve this problem, we suggest a closeness centrality measure that integrates line values and compare it to Freeman's closeness. Small exemplary networks will demonstrate the characteristics of the weighted closeness. The Island approach is used to show the effectiveness of the weighted closeness compared to other centrality measures. Finally, the weighted closeness will be tested on a real-world telecommunications graph provided by a large Austrian mobile service provider.

Using Latent Class Analysis To Categorize Patterns of Psychiatric Disorders: The Case of Nicotine Withdrawal

✱Tania Robbins, RTI International, *trobbins@rti.org*; Scott Novak, RTI International; Sara Calvin, RTI International

Key Words: latent class analysis, nicotine withdrawal

Nicotine dependence has been shown to be consistently correlated with difficulty in quitting smoking and the unpleasant effects of withdrawal. Reducing the severity of withdrawal symptoms has been a focus of treatment tools in an effort to reduce the likelihood of relapse among those quitting smoking. This paper discusses a methodology for determining the distribution of withdrawal subtypes in a population of lifetime cigarette smokers and modeling the association between nicotine withdrawal symptoms and persistent smoking. A latent class analysis plus a logistic regression model were used to determine the most clinically relevant subtypes of nicotine withdrawal symptoms and to predict past year smoking as a function of withdrawal symptoms and psychiatric measures. The methodology is illustrated using a sample from the National Epidemiologic Survey on Alcohol and Related Conditions.

Reading Literacy Trajectories Among Different Gender and Ethnicity

✱Man Hung, The University of Utah, 161 University Village, Salt Lake City, UT 84108, *mh1238@hotmail.com*

Key Words: reading achievement, ECLS-K, ethnicity, growth model, elementary education, gender

There were two main goals of this study: (1) to assess reading performance among students of different gender and ethnicity, and (2) to look at the change of their reading performance across time. The data for this study came from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), K-5th Grade. Using multilevel modeling, various linear growth models were tested to examine the trends in reading literacy. Results indicate that girls outperformed boys throughout the years, but they had the same growth rate as boys starting from third grade. White students had consistently outperformed Black and Hispanic students across time. Asian students slightly outperformed White students for the first two years of elementary school; but this gap had closed in later years. However, the gap between high and low performing students became larger and larger as time progressed.

Variable Selection for Propensity Score Models

✱Bing Yu, University of Toronto, 30 Charles St. West, 1407, Toronto, ON M4Y 1R5 Canada, *bingyu@oise.utoronto.ca*; Guanglei Hong, University of Toronto

Key Words: selection bias, potential outcomes, efficiency, interaction

Propensity score method has been increasingly used in social scientific and educational research. Some existent evidence suggests that selecting variables on the basis of their associations with the outcome is preferable to variable selection on the basis of their associations with the treatment. An important class of variables, however, has been overlooked in the literature. These are variables that do not show main effects on the outcome yet their associations with the potential outcome under one treatment are different from their associations with the potential outcome under an alternative treatment. Through mathematical derivation and simulation, we find that failure to include such variables in a propensity score model will introduce bias in the estimated average treatment effect and at the same time may increase variance and mean square error of the estimation.

Discrepancies in Cohen's Proposed Effect Sizes

✱James Schmeidler, Mount Sinai School of Medicine, Department of Psychiatry–Box 1230, 1 Gustave L. Levy Place, New York, NY 10029, *James.Schmeidler@mssm.edu*

Key Words: power analysis, effect sizes

In "Statistical Power Analysis for the Behavioral Sciences," Cohen proposed "small," "medium," and "large" effect size values for a variety of analyses, and showed comparability for some analyses. However, for two analyses testing the difference of a binomial proportion from .5 with very similar power, his proposed proportions are discrepant, .55, .65, and .75 from .60, .74, and .86. Cohen proposed identical effect size values for one- and two-sample versions of any analysis. In a matched-sample design with unrelated samples, one- and two-sample t-tests have the same power analysis but different effect sizes and algorithms, so power differs for his proposed values. Consistent larger proposed values for one- than two-sample tests would eliminate the discrepancy. Following Cohen's implicit derivation of proposed effect sizes for other analyses would reduce their discrepancies.

326 Deming Lecture

ASA, ENAR, IMS, SSC, WNAR, Deming Lectureship Committee, Biometrics Section

Tuesday, August 5, 4:00 p.m.–5:50 p.m.

Inference and Improvement in Health Care

✱Donald M. Berwick, Institute for Healthcare Improvement, 20 University Road, 7th Floor, Cambridge, MA 02138, *DBerwick1@IHI.org*

Classical forms of clinical practice evaluation, most importantly the randomized clinical trial, are now widely honored in health care. These forms are the espoused standard in 'evidence-based medicine.' This evidentiary rigor marks great progress from the eras of mere anecdote and habit as the bases for maintaining or introducing a medical procedure. However, the epistemological underpinnings of modern approaches to continual improvement in complex systems in health care appear sometimes to run afoul of these standards of evidence. A broader array of evaluation methods would be helpful and better able to capture and guide learning in complex systems. This lecture will explore the scale and sources of the current apparent conflicts between evidence-based medicine and continual improvement, as well as suggest approaches to their resolution.

327 ASA Presidential Address and Awards

ASA, ENAR, IMS, SSC, WNAR, Biometrics Section

Tuesday, August 5, 8:00 p.m.–9:30 p.m.

Presentations of Awards

✱Mary Ellen Bock, Purdue University, Department of Statistics, West Lafayette, IN 47907-2067, *mbock@stat.purdue.edu*

Certificates of Appreciation for Retiring Editors - Mary Ellen Bock Samuel S. Wilks Memorial Award - Daniel Zelterman, Chair Samuel S. Wilks Memorial Medal Committee Gottfried E. Noether Awards - Carlos Morales, Chair Noether Awards Committee Statistics in Chemistry Award - TBD, Chair SPES Committee on Chemometrics Award of Outstanding Statistical Application - TBD, Chair Award of Outstanding Statistical Application Committee W. J. Youden Award in Interlaboratory Testing - Neil R. Ullman, Chair W. J. Youden Award in Interlaboratory Testing Committee Edward C. Bryant Scholarship Award - Elaine Zanutto, Chair Edward C. Bryant Scholarship Award Committee Gertrude M. Cox Scholarship in Statistics Award - TBD, Chair Subcommittee to Committee on Women in Statistics and the Caucus for Women in Statistics SPAIG Award - Jai Won Choi, Chair Statistical Partnerships Among Academe, Industry, and Government Award Committee

Communicating Statistics and Developing Professionals

✱Peter A. Lachenbruch, Oregon State University, Department of Public Health, Corvallis, OR 97330, *Peter.Lachenbruch@oregonstate.edu*

For many years, statisticians and the ASA have been concerned with communicating our message to appropriate audiences. These include our peers, our clients, and the public. We must know our audience and communicate in their terms by using nontechnical words/pictures for nontechnical audiences; speak to be understood, especially to international groups. A second part of the talk is concerned with our development as scientific professionals. We need to develop ourselves professionally: technical skills are not enough, statisticians must interact with nonstatisticians (Communicating!); we need business skills to develop budgets, to conduct interviews and be interviewed, to know how to hire (there may be legal pitfalls), to learn to negotiate, how to set fees, to manage our time and to organize our offices and lives.

Presentation of Founders Awards and New ASA Fellows

✱Mary Ellen Bock, Purdue University, Department of Statistics, West Lafayette, IN 47907-2067, *mbock@stat.purdue.edu*

Founders Award - Mary Ellen Bock, Chair Founders Award Committee New ASA Fellows - Roderick Joseph Little, Chair Committee on Fellows

328 Section on Quality and Productivity Roundtable with Coffee (fee event)

Section on Quality and Productivity

Wednesday, August 6, 7:00 a.m.–8:15 a.m.

Will Work for Data: Observations on Life as an Industrial Statistician

✱Theresa Utlaut, Intel Corporation, 5200 NE Elam Young Pkwy, RA3-353, Hillsboro, OR 97124, *theresa.l.utlaut@intel.com*

Key Words: industrial statistics, employment, salary, career path

Some people, when thinking of the field of industrial statistics, think of endless days of rehashing statistical process control and design of experiments and little else. Many of those people have likely never worked as an industrial statistician and would be surprised to learn what a career in industrial statistics entails. It can be an exciting career if made into one. This roundtable coffee is intended for graduate students and recent graduates deciding on a career path in statistics. Discussion topics will vary depending on the interests of participants, but can include personal and professional observations from industry, expected salaries, common tasks of industrial statisticians, how to be more employable, and taking control of your career.

329 Section on Statistical Education Roundtable with Coffee (fee event)

Section on Statistical Education

Wednesday, August 6, 7:00 a.m.–8:15 a.m.

Incorporating GAISE in Online Introductory Statistics Courses

✱Sue Schou, Idaho State University, PO Box 4043, Pocatello, ID 83205, *schosue@isu.edu*

Key Words: GAISE, online, active learning, technology

We will discuss how to foster active learning and incorporating innovative technologies in an online setting. The remaining four recommendations of the GAISE college report will be discussed as time permits.

330 Section on Statistics and the Environment Roundtables with Coffee (fee event)

Section on Statistics and the Environment
Wednesday, August 6, 7:00 a.m.–8:15 a.m.

Bayesian Space-Time Disease Mapping: Issues and Opportunities

✱Andrew B. Lawson, University of South Carolina, Dept of Epoidemiology & Biostatistics, 800 Sumter Street, Columbia, SC 29208, *alawson@gwm.sc.edu*

Key Words: Bayesian, spatial, health, pollution, space-time

Bayesian hierarchical modeling of georeferenced health data is often concerned with temporal changes to maps of disease. The focus can be related to estimation of the linkage to environmental covariates or simply the relative risk estimation over time. In this roundtable, issues relating to space-time modeling will be considered in relation to environmental risk gradients. In particular, the concern that space-time interaction effects could play a significant part in both description of the risk surface as it changes and the relation with covariates displaying subtle interactions. An example of this is the possible change in risk relations over time for pollution emission sources and their risk imprint. Discussion will focus on unobserved and observed time-labeled pollution events.

20 Questions Statisticians Should Ask About Climate Change

✱Edward J. Wegman, George Mason University, 4400 University Drive, MS 6A2, Fairfax, VA 22030-4422, *ewegman@gmail.com*; Yasmin H. Said, George Mason University; David W. Scott, Rice University

Key Words: paleoclimate reconstruction, climate modeling, statistical/stochastic issues, IPCC 2007 report, Global Warming

Last fall, the ASA sponsored a workshop, “A Statistical Consensus on Global Warming,” that was held at the National Center for Atmospheric Research. The workshop was aimed at understanding the areas of agreement from a statistical perspective. In light of the issuance of the IPCC 2007 report this year and the general lack of success of the Kyoto Accord to stem greenhouse gas emissions, this workshop has an important role in developing the consensus on statistical issues. Although both paleoclimate reconstruction and climate modeling have many fundamentally statistical/stochastic issues, the convergence of the perspectives of statisticians and climate scientists is not great. We seek to raise some of the statistical issues related to inferences about climate change.

331 Section on Statistics in Epidemiology Roundtable with Coffee (fee event)

Section on Statistics in Epidemiology
Wednesday, August 6, 7:00 a.m.–8:15 a.m.

Methods for Assessing Exposures to Mixtures of Chemicals

✱Amy H. Herring, The University of North Carolina at Chapel Hill, CB #7420, Department of Biostatistics, Chapel Hill, NC 27599, *aherring@bios.unc.edu*; Enrique F. Schisterman, National Institute of Child Health and Human Development; ✱Rajeshwari Sundaram, National Institutes of Health, 6100 Executive Blvd., 7B03Q, Rockville, MD 20852, *sundaramr2@mail.nih.gov*

Key Words: study design, correlated exposure, toxicology, mixtures

While laboratory studies are often designed to evaluate effects of single chemical exposures, most human populations are exposed to mixtures of a variety of known and potential toxins. Study of toxicity of environmental agents whose quantification may be mixtures is complex. Much current work focuses on only a parent compound, a common metabolite, or a summary measure calculated over a class of compounds. Interesting issues for both study design and analysis include disentangling single-chemical versus multiple-chemical effects, measurement error, and limit of detection issues. We will discuss new study design and analysis methods for both laboratory and epidemiological studies that may advance our understanding of health effects of exposures to mixtures of chemicals.

332 Section on Teaching Statistics in the Health Sciences Roundtable with Coffee (fee event)

Section on Teaching Statistics in the Health Sciences
Wednesday, August 6, 7:00 a.m.–8:15 a.m.

Enhancing Statistical Literacy in the Medical Professions

✱Penelope Pekow, University of Massachusetts Amherst, Program in Biostatistics & Epidemiology, Arnold House, Amherst, MA 01003, *ppekow@schoolph.umass.edu*; Carol Bigelow, University of Massachusetts Amherst

Key Words: medical statistics, evidence based medicine, basic statistics

Increasingly sophisticated statistical methods and presentations are common in the most widely read medical journals. While physicians are expected to keep abreast of the latest in evidence-based medicine, training in basic statistics in medical school and residency is minimal, or even absent. Including statistics in a medical education system suffering from real time constraints is difficult. Possible viable approaches for enhancing statistical literacy in this population include, but might not be limited to, journal clubs, web resources, and CME programs.

333 Late-Breaking Session II: What Can Statistical Methods Tell Us About Steroid Use and Its Effects Among Major League Baseball Players?

ASA, ENAR, IMS, SSC, WNAR, Section on Statistics in Sports

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

What Can Statistical Methods Tell Us About Steroid Use and Its Effects Among Major League Baseball Players?

✱James H. Albert, Bowling Green State University, 407 Math Science, Bowling Green, OH , *albert@math.bgsu.edu*; ✱Michael J. Schell, Moffitt Cancer Center, 12902 Magnolia Drive, Tampa, FL 33612-9416, *michael.schell@moffitt.org*; ✱Andrew Dolphin, Steward Observatory, P.O. Box 26732, Tucson, AZ 85726, *adolphin@raytheon.com*; ✱Phil Birnbaum, Society for American Baseball Research, Statistical Newsletter Editor, 88 Westpointe Cres., Nepean, ON K2G 5Y8 Canada, *birnbaum@sympatico.ca*

In this panel discussion, we will give some background material on how one measures player performance in baseball. There are decisions to be made regarding the choice of performance measure and how one properly adjusts a measure for the era of play. A player has an associated career trajectory and associated peak performance and there are different ways of comparing player trajectories. We review results on hitting performance and see how these results can be applied to evaluating pitching performance. Are there unusual patterns of hitting of the "steroid sluggers" such as Barry Bonds and Mark McGwire, and are there similar unusual patterns of pitching of pitchers such as Roger Clemens?

334 Introductory Overview Lecture: Sample Size and Related Issues

ASA, ENAR, IMS, SSC, WNAR, Biopharmaceutical Section, Section on Government Statistics, Section on Survey Research Methods, Section on Teaching Statistics in the Health Sciences, Social Statistics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

What Are the Crucial Error Rates To Consider in Sample-Size Analysis?

✱Ralph O'Brien, Case Western Reserve University, Center for Clinical Investigation - BRB109, 10900 Euclid Avenue, Cleaveland, OH 44106-4961 USA, *obrienralph@gmail.com*

Key Words: power analysis

We begin by reviewing the key concepts in classical power and sample-size analysis for regular (frequentist) hypothesis testing. More deeply, we ask: How should such planning help convince us that a significant p-value indeed reflects a true research hypothesis? By using judgments about the probability that the null hypothesis is false, we apply Bayes' Theorem to assess the probability that a significant p-value will be a Type I error or a nonsignificant p-value will be a Type II error. We dub these the "crucial" Type I and II error rates, and we show that they can differ greatly from their classical counterparts. All ideas are illustrated by examining a small preliminary study that tested a very speculative treatment for atherosclerosis and became enthusiastically reported due to its significant p-value supporting efficacy. Unfortunately, the crucial Type I rate may have been over 85%.

A Good N Despite a Bad Start: A Practical Guide to Easy Internal Pilots

✱Keith E. Muller, University of Florida, Department of Epidemiology and Health Policy Reseach, 1329 SW 16th Street Room 5125, PO Box 100177, Gainesville, FL 32610-0177, *Keith.Muller@biostat.ufl.edu*

Key Words: Design, Sample Size, Power, Small Sample, Adaptive, Linear Model

An internal pilot design provides an easy and low-cost path to a good sample size despite starting with inadequate information. For Gaussian data, an internal pilot uses an interim variance estimate to give a sample size adjustment (without interim data analysis). A brief introduction highlights the practical benefits and limitations of internal pilot designs. In contrast to many types of adaptive designs for sample size re-estimation, internal pilots come from a solid analytic foundation, which allows balanced control of sample size, statistical power, and cost. The real world introduces some barriers that can usually be overcome by simple strategies. Ready to use methods for large and small designs will be described and recommended for study planning, interim sample size adjustment, and final analysis.

335 The Life of a Statistician: In Memory of Professor Jack Lee ▲

Memorial, Section on Nonparametric Statistics, Committee on ASA Archives and Historical Materials

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

The Life of a Statistician: In Memory of Professor Jack Lee

✱Jane-Ling Wang, University of California, Davis, 1 Shields Ave., Department of Statistics, Davis, CA 95616, *wang@wald.ucdavis.edu*

I will present the current status of the International Chinese Statistical Association and Jack Lee's contributions to the association.

A Robust Approach to Joint Modeling of Mean and Scale Covariance for Longitudinal Data

✱Tsung-I Lin, National Chung Hsing University, Department of Applied Mathematics, Taichung, 402 Taiwan, *tilin@amath.nchu.edu.tw*; Yun-Jen Wang, National Chiao Tung University

Key Words: Covariance structure, Maximum likelihood estimates, Reparameterization, Robustness, Outliers, Prediction

We propose a multivariate t regression model with its mean and scale covariance modeled jointly for the analysis of longitudinal data. A modified Cholesky decomposition is adopted to factorize the dependence structure in terms of unconstrained autoregressive and scale innovation parameters. We present three distinct representations of the log-likelihood function of the model and study the associated properties. A computationally efficient Fisher scoring algorithm is developed for carrying out maximum likelihood estimation. The technique for the prediction of future responses in this context is also investigated. The implementation of the proposed methodology is illustrated through two real-life examples.

Nonparametric Monotone Regression for Generalized Linear Models with Applications to Wafer Acceptance Tests

✱Jyh-Jen H. Shiau, National Chiao Tung University, Institute of Statistics, 1001 Ta-Hsueh Road, Hsinchu, Taiwan, 30050 Taiwan, *jyhjen@stat.nctu.edu.tw*

Key Words: Monotone nonparametric regression, generalized linear model, Bernoulli data, natural cubic splines, smoothing splines, Wafer Acceptance Tests

Motivated by an engineering control problem of wafer acceptance tests (WAT) in semiconductor manufacturing, we develop a new nonparametric monotone smoothing-spline-based smoother for analyzing responses from exponential families. The new method modifies the monotone smoother developed by Zhang (2004) and then combines with the methodology developed by Gu (2002) for data from exponential families. An efficient algorithm is provided. A simulation study demonstrates that the proposed method performs well in the regression models with Bernoulli responses. In terms of the averaged squared error, the proposed monotone estimator outperforms the unconstrained smoother when the latter produces non-monotone estimates, while retaining about the same performance otherwise. It is demonstrated that the proposed method can be used in screening WAT test items for more stringent engineering control.

Multidimensional Scaling for Large Genomic Data Sets

Jengnan Tzeng, Academia Sinica; ✱Henry H.S. Lu, National Chiao Tung University, 1101 Ta Hsueh Road, Hsinchu, 30050 Taiwan, *hslu@stat.nctu.edu.tw*; Wen-Hsiung Li, The University of Chicago

Key Words: Multi-dimensional scaling (MDS), dimension reduction, clustering, K-means, microarray, cell cycle

We developed a new rapid metric MDS method with a low computational complexity, making metric MDS applicable for large data sets. Computer simulation showed that the new method of split-and-combine MDS (SC-MDS) is fast, accurate and efficient. Our empirical studies using microarray data on the yeast cell cycle showed that the performance of K-means in the reduced dimensional space is similar to or slightly better than that of K-means in the original space, but about three times faster to obtain the clustering results. Our clustering results using SC-MDS are more stable than those in the original space. Hence, the proposed SC-MDS is useful for analyzing whole genome data.

336 Mixture Models: A Tool for Multilayered Clustering and Dimension Reduction

IMS, WNAR

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

EM-Test for Finite Mixture Models

✱Jiahua Chen, University of British Columbia, Department of Statistics, Vancouver, BC V6T 1Z2 Canada, *jhchen@stat.ubc.ca*

Key Words: Homogeity test, Infinite Fisher information, compactness conditions, limiting distribution

Most existing methods in the literature for testing of homogeneity, explicitly or implicitly, are derived under the condition of finite Fisher information and a compactness assumption on the space of the mixing parameters. The finite Fisher information condition can prevent their usage to many important mixture models, such as the mixture of geometric distributions, exponential distributions and more generally mixture models in scale distribution families. The compactness assumption is relatively harmless, yet it can be awkward to specify a compact region for the mixing parameters in applications. In this presentation, we introduce an EM-test, which is shown to be free of all these shortcomings and have very simple limiting distributions. Current results indicate that this method has the potential to be generally applicable.

Modal Inference and Its Application to High-Dimensional Clustering

✱Surajit Ray, Boston University, 111 Cummington Street RM 222, Boston, MA 02215, *sray@math.bu.edu*

Key Words: Clustering, Modal Clustering, Topography, Modal EM, Multiscale

Ray and Lindsay (2005) show that the topography of multivariate mixtures, in the sense of their key features as a density, can be analyzed rigorously in lower dimensions by use of a ridgeline manifold that contains all critical points as well as the ridges of the density. To use this rich feature for data analysis we first construct an extension of EM algorithm that can be used to find the modes of a mixture density. Even in very high dimensions the computational complexity of our EM algorithm is extremely low. These tools can be used in various ways. For one, we can take a conventional mixture analysis and cluster together those components whose contribution is actually unimodal. We can also turn kernel density estimation into clustering tool in which the data points become identified with each other by their association with a common mode of the density estimator.

Mixture Models for Document Clustering

✱Edward J. Wegman, George Mason University, 4400 University Drive, MS 6A2, Fairfax, VA 22030-4422, *ewegman@gmail.com*

Automatic clustering and classification of documents within corpora is a challenging task. Often, comparing word usage within the corpus, the so-called bag-of-words methodology, does this. In this talk, in addition to comparing word usage, we extract additional endogenous features of the documents and use a mixture model density estimate to localize the clusters.

337 Privacy Breaches in Federal Data Collections ▲

Section on Survey Research Methods, Section on Government Statistics, Section on Statisticians in Defense and National Security

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Applying Federal Data Security Policy to Statistical Agency Practices

✱Rochelle W. Martinez, Office of Management and Budget, 304 11th St, SE, Washington, DC 20003, *rmartinez@omb.eop.gov*; John W. Barkhamer, Office of Management and Budget

Key Words: data breach, confidentiality, security, personally identifiable information, data stewardship, social security number

Many in the media dubbed both 2005 and 2006 the “Year of the Data Breach,” given well-publicized breaches across the private sector, academia and government. The Office of Management and Budget (OMB) led federal efforts with the issuance of a key policy memorandum addressing data breach notification, unnecessary collection and retention of personally identifiable information (PII) (e.g., social security numbers), and a number of other privacy and security aspects. While government-wide in scope, the guidance understandably has a distinct effect on statistical agencies whose missions often involve collecting PII and whose success rests on maintaining the public's trust. This paper focuses on OMB requirements as they affect statistical agency programs, and suggests ways that agencies can move beyond simple compliance to a strategic approach to data stewardship.

U.S. Census Bureau's Approach To Address OMB's Guidelines and Rules Relating to Privacy Breaches in Federal Statistical Agencies

✱Nancy M. Gordon, U.S. Census Bureau, HQ 8H128, Washington, DC 20233-0001, *nancy.m.gordon@census.gov*

Upon the release of the Identify Theft Task Force's September 2006 memorandum, the Census Bureau recognized the need to strengthen its processes to address potential data breaches. Both the report of the Identify Theft Task Force and the draft OMB memorandum were used as the foundation for the enhanced policy. The bureau decided to use a widely accepted model of developing a risk score based on the likelihood of the event occurring and the impact of the event, an approach that was included in the OMB memorandum. The bureau incorporated an existing internal board into the revised process, the Data Stewardship Executive Policy Committee (DSEP). The DSEP reviewed the policy, and it was adopted on December 15, 2006. This paper will describe the Census Bureau's processes and its experiences since they were adopted.

The Impact of a Privacy Breach on Survey Participation in a National Longitudinal Survey

✱Marilyn Seastrom, National Center for Education Statistics, 1990 K Street, NW, Washington, DC 20006, *Marilyn.Seastrom@ed.gov*; Jennifer Park, National Center for Education Statistics

Key Words: Confidentiality, Data breach, Disclosure risk, Survey participation, Response rates

A longitudinal study with four rounds of data collection planned over a 5-year period experienced a data loss during field operations in the second round of data collection. Cases with data loss were evaluated for risk of disclosure. All cases with data loss were notified, and the level of potential risk was reported to each case. Those with a potential risk of disclosure were offered a year of credit monitoring. Cases with data loss were allowed to opt out of the data collection and their data were removed from the computer files. The disclosure risk analysis is described; the numbers of cases impacted are reported, along with the number of cases that opted out during the second round of data collection; and the participation of the impacted cases compared to the rest of the sample is compared in round three of the data collection.

338 Stirring the Pot: Radical Ideas in Statistics Education

Section on Statistical Education

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Statistical Thinking: A New First Course in Statistics

✱Bruce Dunham, University of British Columbia, 333 - 6356 Agricultural Way, LSK Building, Vancouver, BC V6T 1A4 Canada, *b.dunham@stat.ubc.ca*

Key Words: statistical education, introductory statistics, course design

A model is presented for an introductory course in Statistics that can be taken before, after or alongside a more traditional first course in the discipline. This year UBC introduced "Statistical Thinking," a modular course available to all students with grade 12 Mathematics. The course aims to instill statistical literacy in the learners, while also exploring interesting applications of the field in modern research and society in general. Here the design, content and practicalities of the new course are described.

It's Time To Retire 'N >= 30'

✱Tim Hesterberg, Google, Inc, 651 N. 34th Street, Seattle, WA 98103-8846, *timhesterberg@gmail.com*

Key Words: Central limit theorem, bootstrap, normal distribution, diagnostics, resampling

The old rule of using z or t tests or confidence intervals if n >= 30 is a relic of the pre-computer era, and should be discarded in favor of bootstrap-based diagnostics. These results will surprise many statisticians, who don't realize how lousy the classical inferences are. For example, 95% confidence intervals should miss 2.5% on each side, and we might expect the actual noncoverage to be at most 10% off (2.25%-2.75%). Using a t interval, this requires n > 5000 for a moderately skewed (exponential) population! There are better confidence intervals and tests, bootstrap or otherwise. The bootstrap also offers pedagogical benefits in teaching sampling distributions and other statistical concepts, offering actual distributions that can be viewed using histograms and other familiar techniques.

Concepts of Statistical Inference: A Randomization-Based Curriculum

✱Allan Rossman, California Polytechnic State University, Dept of Statistics, San Luis Obispo, CA 93407, *arossman@calpoly.edu*; Beth Chance, California Polytechnic State University; George Cobb, Mt. Holyoke College; John Holcomb, Cleveland State University

Key Words: statistics education, conceptual understanding, randomization test, simulation

We present an overview of a project that aims to transform the introductory statistics course by putting the process of statistical investigations and the logic of statistical inference at its center. One key idea is to present the entire process of statistical investigations (from data collection to data analysis to statistical inference to drawing conclusions) over and over throughout the course. Our hope is that students develop a deeper understanding of key concepts as they encounter them multiple times in new situations. A second key idea is to use simulation of randomization tests and sampling distributions, rather than normal-based procedures, as the starting point for student investigations of the concepts of inference. We provide examples of student activities and also present preliminary findings from our work on assessing student learning with this approach.

Guided Interdisciplinary Research Projects

✱Shonda Kuiper, Grinnell College, IA 50112, *kuipers@grinnell.edu*

Key Words: education, interdisciplinary, technology

This presentation discusses a course comprised entirely of guided interdisciplinary projects. Each project starts with a multiday lab module that introduces a relatively advanced undergraduate statistical technique that is becoming commonly used in other disciplines. Following the introductory lab, students experience the role of a research scientist with step-by-step instructions that guide them through reading and evaluating primary literature in multiple disciplines, preparing a proposal for analysis, planning and carrying out experiments, and presenting the results. To introduce students early in their college careers to the broad applicability and exciting career opportunities is statistics, a high-school AP statistics course is the only prerequisite for this course.

339 Statistical Consulting and Collaboration in Private Industries ●▲

Section on Statistical Consulting

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Statistical Collaboration at Boeing: The Good, the Better, and the Best

✱Sabyasachi Basu, The Boeing Company, 9702 173rd CT NE, Redmond, WA 98052, *sabyasachi.basu@boeing.com*

Key Words: Consulting, Collaboration, Engineering, Communication

In this presentation, we will talk about the role of a statistician within the Boeing Company. We work with colleagues with expertise in different fields such as mathematics, engineering, medical, marketing, and product development. Not only do we have to learn the language of our colleagues, we also have to be able to communicate statistical ideas and results. The work is challenging and rewarding at the same time and we will talk about some of these challenges and rewards.

Statistical Consulting in Pharmaceutical Development: What Turns Outsourcing into Collaboration?

✱Russell Helms, Rho, Inc., 6330 Quadrangle Drive, Suite 500, Chapel Hill, NC 27517, *Russ_Helms@RhoWorld.com*

Key Words: consulting, statistical outsourcing, collaboration

Most pharmaceutical and biotech companies outsource at least some statistical work, often to Contract Research Organizations, with mixed success. Companies have a variety of goals and approaches to statistical outsourcing. Some seek specific expertise, such as experience interacting with the FDA, while others outsource only the simplest work to fill in resource gaps. What is the best approach? This paper argues that finding the right partner, then seeking a collaborative relationship, where both the pharmaceutical company and the CRO contribute intellectually, makes the best use of many minds and varied experiences to produce a superior product. It also discusses factors that facilitate collaboration and getting the most from a relationship, and describes steps pharmaceutical companies and CROs can take to enhance collaboration throughout the entire life cycle of a development project.

Statistical Consulting in the Automotive Industry: Opportunities, Challenges, and Impact

✱Lonnie Vance, General Motors (Retired), 6500 Red Oak Drive, Troy, MI 48098, *LCV3@dcx.com*

Key Words: statistical consulting, automotive industry, challenges

Many opportunities exist for statistical consulting in the automotive industry. These include one-on-one consulting with scientists and engineers, assisting green belts and black belts with Six Sigma projects, and consulting and collaboration with others at department, division, or corporate levels. The challenges are many. These range from too little time to do analyses, lack of appropriate data, unable to publish the case study/results, communication issues, following ethical guidelines for statistical practice, and management not liking the answers. As Deming told us, the impact may be “unknown and unknowable.” If possible, quantify the estimated cost savings with dollars.

340 Innovative Methods for Imputation ●

Section on Health Policy Statistics, Section on Government Statistics, Section on Survey Research Methods

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Evaluation of Multiple Imputation by Ordered Monotone Blocks in an Anthrax Vaccine Trial

Michela Baccini, University of Florence; Constantine E. Frangakis, Johns Hopkins University; Fan Li, Harvard Medical School; Fabrizia Mealli, University of Florence; Brian D. Plikaytis, Centers for Disease Control and Prevention; Charles E. Rose, Centers for Disease Control and Prevention; Donald B. Rubin, Harvard University; ✱Elizabeth R. Zell, Centers for Disease Control and Prevention, 1600 Clifton Road, MS-C09, Atlanta, GA 30329, *ezell@cdc.gov*

Key Words: Multiple Imputation, Randomized Control Trial, Missing Data, Evaluation, Clinical trial

In randomized trials that require repeated visits, it is inevitable that some people will miss some visits. Data should be collected despite these protocol violations, but it is difficult to obtain immunogenicity or reactogenicity data for missed visits. Some sort of imputation is required if ITT analyses are to be done using all randomized subjects. Multiple imputation has been shown to be valid under a broad range of circumstances. CDC’s current Anthrax Vaccine Trial presents challenges because of the ~2000 measurements per person and the ~200 subjects in each treatment arm. We employed a novel ordered monotone-blocks multiple-imputation method. The evaluation involved creating 200 datasets with realistic missing data patterns for 4 immunogenicity and 28 reactogenicity measures. Here we present the results from this evaluation of the validity of the multiple imputations in this trial.

Evaluation of Imputation of Covariates in an Impact Analysis with Regression Adjustment

✱Eric Grau, Mathematica Policy Research, Inc., 600 Alexander Park, Princeton, NJ 08543, *egrau@mathematica-mpr.com*; Sue Ahmed, Mathematica Policy Research, Inc.

Key Words: imputation, sequential hot deck, impact analysis, random assignment, regression adjustment

In an impact analysis using random assignment, researchers often deal with missing values in both the covariates and the outcome variables of regression models. Clearly rigorous methods are needed to impute missing values in the outcome variables to minimize the potential bias in impact assessments. When imputation is applied to covariates of the regression analyses, the effect of imputation is less clear on impact analyses. This paper assesses this effect, using a random assignment evaluation of the Growing America Through Entrepreneurship (GATE) program. Two outcome variables used in the original evaluation are modeled against a set of 10 covariates, a treatment indicator, and variables associated with the site of the evaluation. Impacts are assessed with different levels and types of missingness in the covariates with values imputed using a sequential hot deck.

Multiple Imputation for Protecting Data Confidentiality: Applications by the German Institute for Employment Research

Jerome Reiter, Duke University; ✱Joerg Drechsler, German Institute for Employment Research, IAB, Regensburger Strabe 104, Nurnberg, 90478 Germany, *Joerg.Drechsler@iab.de*; Susanne Rässler, Otto-Friedrich University Bamberg

Key Words: Confidentiality, Disclosure, Imputation, Public use

The German Institute for Employment Research (IAB) is using multiple imputation to generate a public use file of the IAB Establishment survey. The imputations serve to protect data confidentiality while enabling valid inferences for wide classes of estimates. In this talk, we describe some of our imputation strategies, focusing on the tradeoffs among data quality, disclosure risks, and computational feasibility.

Heavy Children in Motor Vehicle Crashes: Imputation for Covariates with Missingness and Measurement Errors

✱Dawei Xie, University of Pennsylvania, 617 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, *dxie@mail.med.upenn.edu*; Yulei He, Harvard Medical School

Key Words: AIS 2 injury, body mass index, missing covariate, multiple imputation, NHANES

It is important to evaluate the efficacy of child restraints for protecting children of different sizes in crashes. Child injury status and parent-reported height and weight were obtained from a complex survey, and a logistic regression model is fitted. However, ignoring the misreporting and missing values in height and weight might lead to bias in model estimates. We propose to account for the missingness and reporting error using a two-stage multiple imputation approach. In the first stage, incomplete parent-reported height and weight are imputed. In the second stage, the true height and weight are imputed using a measurement error model established from ancillary data. Complex survey design is taken into account in both stages of imputation as well as the analysis model.

341 Nested and Crossed Random Effects in Nonlinear Models ●

Biopharmaceutical Section, Section on Bayesian Statistical Science, Section on Health Policy Statistics, Biometrics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Challenges and Opportunities for Nonlinear Mixed Models in Biological Assays

✱David Lansky, Precision Bioassay, 257 S. Union St., Burlington, VT 05401-5428, *david@precisionbioassay.com*

Key Words: bioassay, model selection, variance components, mixed models, crossed error

Modern biological assays contain many sources of variation due to complex design structures. Randomized incomplete-block strip-plots protect assays from many sources of bias, but are not practical without robots. Related manual non-randomized designs protect against up to quadratic location bias. Analysis models contain more random effects than can be estimated; model selection is important. Equivalence testing for fixed effects selection is appropriate and leads to questions whether AIC & BIC (used in selecting random effects) are analogous to difference rather than equivalence testing. Assays designed for and fit with mixed models handle non-additive effects particularly well, yielding large improvements in assay performance. Statistical issues include: how to analyze a non-randomized design and whether equivalence-based methods are more appropriate for selection of random effects.

Semiparametric Bayesian Approaches to Nonlinear Mixed Effects Models

✱Peter Mueller, M.D. Anderson Cancer Center, 1515 Holcombe Blvd., Unit 447, Houston, TX 77030, *pmueller@mdanderson.org*

Key Words: nonparametric, Dirichlet process

In this talk we review the nonparametric Bayesian approach to inference for unknown functions, in particular for unknown mean functions and random effects distributions in nonlinear mixed models. The paradigm of nonparametric Bayesian inference is to recognize the unknown function or density as an unknown quantity and proceed with a probability model on the appropriate function space. We use three examples to illustrate the use of popular Dirichlet process models.

Bayesian Models for Repeatedly Repeated Data

✱Gary L. Rosner, University Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 447, Houston, TX 77030, *glrosner@mdanderson.org*; Peter Mueller, M.D. Anderson Cancer Center; Fernando A. Quintana, Pontifícia Universidad Católica de Chile

Key Words: Bayesian inference, Nonparametric models, Loss of heterozygosity, Repeated measures, Binary sequences, Hierarchical models

We discuss inference for data with repeated measurements at many levels. Inference typically concerns characteristics of the repeated measurements within repeating cycles or dependence across cycles or both. We illustrate modeling and inference with an example relating to alterations of patients' DNA represented by sequences of indicators of loss of heterozygosity. The data involve three nested levels of repetition for each patient: chromosomes, regions within chromosomes, and single nucleotide polymorphisms nested within regions. Our Bayesian semiparametric hierarchical model for these multi-level repeated binary data includes a mixture-of-Markov-chains model, defined with respect to the Markov transition probabilities and a nonparametric prior for the random mixing measure.

342 Statistical Challenges in Vaccine Clinical Studies ●▲

ENAR, Biopharmaceutical Section, WNAR, Section on Health Policy Statistics, Biometrics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Issues and Challenges in Assessing Vaccine Efficacy and Correlates of Protection

✱Ivan S.F. Chan, Merck Research Laboratories, 351 N. Sumneytown Pike, UG1CD-38, North Wales, PA 19454, *Ivan_Chan@Merck.Com*

Key Words: Vaccines, Biomarkers, Efficacy, Sorrogate Endpoints

Vaccines are typically developed for disease prevention in healthy subjects. The unique characteristics of vaccine development pose special challenges in designing vaccine efficacy trials. In addition, immune responses are measured to identify potential markers that correlate with efficacy. In this talk, we will discuss the statistical issues involved in design and analysis of vaccine efficacy trials and in assessing correlates of protection. We will illustrate some potential difficulties in applying classical methods (such as Prentice's criteria) for validating surrogate endpoints. Then, we will present some approaches in evaluating correlates of protection, including the concept of protective level and the use of statistical models to examine the relationship between efficacy and the whole distribution of immune responses. Real vaccine examples will be used to illustrate the methods.

Assessing How Vaccine Efficacy Depends on HIV Genetics by Competing Risks Failure Time Methods

✱Peter B. Gilbert, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, PO Box 19024, Seattle, WA 98105, *pgilbert@scharp.org*; Yanqing Sun, The University of North Carolina, Charlotte; Ian W. McKeague, Columbia University

Key Words: clinical trial, failure time data, genetics, proportional hazards, semiparametric model, survival analysis

For time-to-event data with finitely many competing risks, the proportional hazards model has been a popular tool for relating the cause-specific outcomes to covariates (Prentice et al., 1978). We extend this approach to allow a continuum of competing risks, in which the cause of failure is replaced by a continuous mark only observed at the failure time. We develop inference for this proportional hazards model in which the regression parameters depend nonparametrically on the mark. This work is motivated by the need to assess how HIV vaccine efficacy depends on the genetic divergence of infecting HIV viruses in trial participants from the HIV strains that are contained in the vaccine. The new approach is evaluated in simulations and applied to a randomized and blinded HIV vaccine trial. Elaborations for discussion include handling missing marks and multiple marks, and model diagnostics.

Establishing Vaccine Safety in Studies with Rare Events

✱Valerii V. Fedorov, GlaxoSmithKline, 1250 S. Collegeville Rd, Collegeville, 19426, *valeri.v.fedorov@gsk.com*

Key Words: vaccine safety study, rare events, early stopping

The primary goal of a vaccine safety trial is to assure that the vaccine is safe for the targeted population. Often that is done by rejecting the null hypothesis that the relative risk of an adverse event attributable to the new vaccine is above a given value, greater than one. The exact probability of type I error for the likelihood score test is evaluated by the straightforward enumeration of the binomial outcomes. The latter shows that the traditionally used normal approximation leads to sample sizes with type I error exceeding the nominal level. For rare adverse events, we recommend the Poisson approximation as a rather simple alternative and develop the conditional and unconditional tests. We also propose optimal randomization strategies which either minimize the total number of adverse cases or minimize the expected number of subjects when the vaccine is unsafe.

Statistical Challenges in Vaccine Clinical Studies: An FDA Perspective

✱Amelia D. Horne, U.S. Food and Drug Administration, 1401 Rockville Pike, HFM 217, Center for Biologics Evaluation & Research, Rockville, MD 20852, *amelia.horne@fda.hhs.gov*

Key Words: efficacy, safety, immunogenicity, clinical trials

This presentation will summarize some of the major statistical issues the FDA must consider in the overall evaluation of preventive vaccines. These issues relate to study design, conduct, and analysis. Ultimately, the FDA must make regulatory decisions regarding licensure and label indications with respect to efficacy, immunogenicity, and safety of the vaccines that are under review. Current issues of interest in evaluating these vaccines will be discussed.

343 Best Practices in Collecting Survey Data on Sexual Orientation

Committee on Gay and Lesbian Concerns in Statistics, Section on Survey Research Methods, Social Statistics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Best Practices in Collecting Survey Data on Sexual Orientation

✱Christopher Carpenter, University of California, Irvine, The Paul Merage School of Business, Irvine, CA 92697-3125, *kittc@uci.edu*; ✱Gary Gates, Williams Institute, UCLA School of Law, Box 951476, Los Angeles, CA 90095-1476, *GATES@law.ucla.edu*; ✱Larry Bye, Field Research Corporation, 601 California Street, San Francisco, CA 94108, *larryb@field.com*; ✱Jana Asher, Carnegie Mellon University, 1111 Jackson Avenue, Takoma Park, MD 20912 USA, *jana@asher-resnick.us*

Key Words: Sexual orientation, Survey design, Response rate, Survey mode, Cognative study, Sample design

While collecting data on sexual orientation, sexual attraction and sexual behavior has become routine for some research purposes, such as studies of STDs, several studies, such as the Institute of Medicine report on lesbian health, have recommended broader data collection in order to enrich the analytical possibilities of a data set. Because sexuality and sexual behavior are often multifaceted, the specific measure of such attributes needs to be closely attuned to the research question. The panel will discuss when it is appropriate to collect data on sexual orientation or behavior, how to sample and contact potential respondents, what to ask, how to design a survey instrument, what mode of data collection is most effective, and what effect asking these questions has on response rates, drawing extensively on cognitive studies and experiences gained through state and national surveys.

344 Using Innovative Sample Design Approaches, Contact Methodologies, and User Input To Reduce Coverage Bias in Surveying Hard-to-Locate Populations ●

Section on Government Statistics, Section on Survey Research Methods

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Developing a Sampling Frame for Postdoctoral Researchers

✱Emilda B. Rivers, National Science Foundation, 4201 Wilson Blvd, Arlington, VA 20223, *erivers@nsf.gov*

Key Words: : Postdoctoral researchers, sampling frame, doctorates

Recently, there have been increases in the number of postdoctoral researchers (postdocs), the duration of postdoc appointments, and the number of non-US postdocs. Understanding the impact of these increases on the US research enterprise and employment markets requires information that is currently unavailable from survey data or other sources. The National Science Foundation (NSF) is engaged in a multiyear project to determine the feasibility of obtaining national-level postdoc statistics for US and non-US doctoral degree holders across all sectors that employ postdocs. A formidable project challenge is frame development and testing methodologies that obtain needed postdoc data from the nonacademic sector and for postdocs with doctorates from foreign universities. This paper will describe the alternative frame development approaches that NSF is exploring to gather postdoc information.

Assessment of Lists for Building a Sample Frame of Academic Postdoc Employers

✱Tina Mainieri, Survey Sciences Group LLC, 220 East Huron Street, Suite 440, Ann Arbor, MI 48104, *tmainieri@surveysciences.com*; Eduardo Galvan, Survey Sciences Group LLC; David Roe, Survey Sciences Group LLC; Emilda B. Rivers, National Science Foundation

Key Words: sample frames, list assessments, postdoctoral researchers

Current national-level statistics for postdoctoral researchers in the United States maybe flawed due to coverage gaps in sampling frames. To begin to rectify this situation, the National Science Foundation is experimenting with methodologies to build a comprehensive sample frame of academic and nonacademic institutions where postdocs work. The methodologies are designed to minimize building an inadequate sample frame that adversely impacts coverage. The multi-step frame building process involves both qualitative and quantitative methods that include administrative list reviews and rapid turnaround surveys. Key to the effectiveness of this process is determining each list's contribution before merging it into a comprehensive sample frame. This paper will discuss the outcomes of the multistage process for the academic institutions component of the comprehensive sample frame for postdocs.

Frame Improvements for the 2007 Commodity Flow Survey

✱Ruth E. Detlefsen, U.S. Census Bureau, SSSD, 4600 Silver Hill Road, Washington, DC 20233, *Ruth.E.Detlefsen@census.gov*; Yukiko T. Ellis, U.S. Census Bureau

Key Words: Frame, Commodity Flow

The Commodity Flow Survey (CFS) is conducted approximately once every five years to provide data on the movement of goods in the United States. Information on the frame available for CFS was inadequate for determining shipper status or developing accurate measures of shipping activity for auxiliaries (establishments of large companies that provide support to the rest of the company). For the 2007 survey, steps were taken to address this issue. This paper will describe the problem and its effect on the 2002 CFS, discuss the new approach and its implications for the frame and subsequent sample design for 2007, and provide an evaluation of the approach.

Getting an Establishment Survey to the Right Person Within the Organization

✱Jeri M. Mulrow, National Science Foundation, 4201 Wilson Blvd. Ste. 965, Arlington, VA 22230, *jmulrow@nsf.gov*; Ray Wolfe, National Science Foundation; Brandon Shackelford, Twin Ravens Consulting

Key Words: establishment survey, contact strategies, different respondents

The National Science Foundation (NSF) has been collecting data on research and development since 1953. In 2004, the NSF embarked upon a series of activities designed to evaluate and renew this important survey. As a result, the entire structure of the survey is being changed. The survey will consist of four distinct parts each focused on different types of data. The four components are aimed at collecting data on (1) financial aspects of the R&D, (2) workforce personnel involved in the R&D, (3) technical aspects of the R&D, and (4) intellectual property issues surrounding the R&D. Different people within the business will need to answer different parts of the survey. This paper discusses the development of the four part questionnaire and the contact strategies needed to get the right component to the right person within the organization.

345 Advances in Bayesian Methods for Health Care Applications ●

Section on Bayesian Statistical Science, Section on Health Policy Statistics

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Determining Latent Dimensions of Disability: Use of Bayesian Methods To Combine Theoretical and Empirical Arguments

✱Mary Beth Landrum, Harvard Medical School, Department of Health Care Policy, 180 Longwood Ave, Boston, MA 02115, *landrum@hcp.med.harvard.edu*

Key Words: disability, Bayesian analysis, Factor analysis

Recent research has demonstrated declines in the prevalence of disability among the elderly population in the United States. Past work aiming to understand causes of disability has relied on simple binary indicators to describe multiple self-reported measures of functioning. Theoretical arguments suggest a progression from physical limitations, to needs for assistance with routine tasks (such as grocery shopping or managing money), and then finally to needs for assistance with the most basic personal tasks such as toileting and feeding. In this talk, we describe the use of latent variable models to summarize disability as measured by a series of individual items of functioning across a smaller number of underlying dimensions. We use theoretical beliefs about the disablement process to develop priors for the number of latent factors and factor loadings of individual items.

Bayesian Model Checking for Multivariate Data

✱Catherine M. Crespi, University of California, Los Angeles, School of Public Health and Jonsson Comprehensive Cancer Center, Center for Health Sciences 51-254, Los Angeles, CA 90095-1772, *ccrespi@ucla.edu*; W. John Boscardin, University of California, Los Angeles

Key Words: Bayesian analysis, Dissimilarity measures, Multivariate data, Posterior predictive model checking

Bayesian models are increasingly used to analyze complex multivariate data. However, methods for checking such models have not been well-developed. We present a method of evaluating the fit of Bayesian models for multivariate data based on posterior predictive model checking (PPMC), a technique in which observed data are compared to replicated data generated from model predictions. We introduce the use of dissimilarity measures for PPMC for multivariate data. This method has the advantage of checking the fit of the model to the complete data vectors or vector summaries with reduced dimension, providing a overall picture of model fit. An application with longitudinal binary data from a clinical trial illustrates the methods.

Bayesian Analysis Using Priors from Historical Data

✱Brian Neelon, Harvard Medical School, 180 Longwood Ave, Boston, MA 02115, *neelon@hcp.med.harvard.edu*; A. James O'Malley, Harvard Medical School

Key Words: Bayesian Hierarchical Model, Evidence Synthesis, Historical Data, Power Prior

We describe a Bayesian approach to incorporating historical data into a current analysis when there is uncertainty about the similarity between the current and historical studies. First, we model the current and historical data under an exchangeability assumption, in which the covariate effects are allowed to differ across studies but are assumed to arise from a common prior distribution. Next, we downweight the impact of the historical data using a power prior distribution (Ibrahim and Chen, 2000). The power prior can prove useful if one believes the exchangeability assumption is itself too strong. We discuss whether the data can guide how much weight to assign to the historical data, or whether this impact factor needs to be user-specified. We apply the method to a cohort study that sought to improve the quality of care in pediatric clinics.

Correcting for Measurement Error in Diagnoses of Post-Traumatic Stress Disorder

✱Juned Siddique, The University of Chicago, Department of Health Studies, 5841 S Maryland Ave., MC2007, Chicago, IL 60637, *siddique@uchicago.edu*; Robert Gibbons, University of Illinois at Chicago; Bonnie Green, Georgetown University Medical School

Key Words: imputation, measurement error, post-traumatic stress disorder

An accurate diagnosis of Post-Traumatic Stress Disorder (PTSD) requires an expert clinician who is able to obtain a detailed trauma history of the patient and then perform a fine-grained analysis of symptom severity. In mental health research it is often unfeasible for participants to be assessed by expert clinicians. Instead, studies must use lay interviewers or rely on questionnaires that measure self-reported symptoms. Both approaches tend to focus on the number of PTSD symptoms rather than symptom severity and can result in a high number of false positives. We present a multiple imputation approach to correct for measurement error in diagnoses of PTSD using a Bayesian ordinal probit model. The method is applied to a depression treatment study where nurse practitioners were twice as likely as psychologists to diagnose PTSD.

Combining Information on Adjuvant Cancer Therapies from Multiple Sources

✱Yulei He, Harvard Medical School, Department of Health Care Policy, 180 Longwood Ave., Boston, MA 02115, *he@hcp.med.harvard.edu*; Alan Zaslavsky, Harvard University Medical School

Key Words: Adjuvant therapy, Cancer, Bayesian Hierarchical model, Misreporting, Multiple imputation, Multivariate outcome

Measures of receipt of adjuvant therapies (chemo- and radiation therapy) are important variables in studies of patterns of cancer care, but no single source reports these therapies with complete accuracy. The Cancer Care Outcomes Research and Surveillance (CanCORS) study collected these variables from multiple sources including a patient survey, medical records abstraction, cancer registry databases, and Medicare claims data. The sample coverages of these sources are different, and the corresponding reporting patterns vary with different accuracy. We propose to impute the true therapy status using Bayesian hierarchical models, synthesizing data from all sources while correcting the misreporting and incorporating the correlations among them. Valid analysis of the therapy variables can then be based on multiply imputed data. We present results using the CanCORS data.

346 Combining Observations and Atmospheric Models To Predict and Understand Weather ●▲

Section on Statistics and the Environment
Wednesday, August 6, 8:30 a.m.–10:20 a.m.

NCAR's Data Assimilation Research Testbed

Jeffrey Anderson, National Center for Atmospheric Research; ✱Timothy Hoar, National Center for Atmospheric Research, 1850 Table Mesa Dr., Boulder, CO 80305, *thoar@ucar.edu*; Nancy Collins, National Center for Atmospheric Research

Key Words: Data Assimilation, adaptive algorithms, atmospheric modeling, education

The Data Assimilation Research Testbed (DART) is an open-source community facility for ensemble data assimilation research, development, and education. DART has a wide range of novel algorithms that deal with commonly encountered errors—a Bayesian algorithm that recommends a multivariate spatially anisotropic localization, for example. In addition, DART supports more novel DA applications like parameter estimation, sensitivity analysis, observing system design, and smoothing. The extensive data assimilation tutorial is appropriate for undergraduate and graduate instruction and is designed to be run with low-order models included with DART. Results from assimilations of millions of observations with state-of-the art climate models and regional weather forecast models are presented, as is a discussion on areas ripe for contributions from statisticians.

Experimental Implementation of an Ensemble Adjustment Filter for an Intermediate ENSO Model

✱Alicia Karspeck, National Center for Atmospheric Research, 1850 Table Mesa Dr., Louisville, CO 80305, *aliciak@ucar.edu*; Jeffrey Anderson, National Center for Atmospheric Research

Key Words: data assimilation, Kalman filter, ENSO, covariance localization

The assimilation of sea surface temperature (SST) anomalies into a coupled ocean-atmosphere model of the tropical Pacific is investigated using an ensemble adjustment Kalman filter (EAKF). The coupled numerical model used is routinely used for simulation and prediction of the El Nino Southern Oscillation. Operating under the "perfect model" experiment paradigm, we investigate how and why changes in the filter parameters (ensemble size, covariance localization, and covariance inflation) affect the quality of the analysis. It is shown that isotropic covariance localization does not benefit the analysis even when a small number of ensemble members are used. These results suggest that destruction of the dynamical balance between variables caused by localization is more detrimental than spurious correlation due to small ensemble size.

The Impacts of Radio Occultation Data on Analyses and Prediction of Tropical Cyclones Using WRF and an Ensemble Forecast System

✱Hui Liu, National Center for Atmospheric Research, 1850 Table Mesa Dr., Boulder, CO 80305, *hliu@ucar.edu*; Chris Snyder, National Center for Atmospheric Research; Jeffrey Anderson, National Center for Atmospheric Research; Y. H. Kuo, National Center for Atmospheric Research

Key Words: data assimilation, weather forecasting, hurricanes, GPS

Radio occultation (RO) refractivity from the COSMIC satellites is available globally since 2006. The data has water vapor and temperature information, is not contaminated by clouds or precipitation, and have relatively high vertical resolution in the lower troposphere. Observations of RO refractivity are assimilated in the WRF/DART ensemble kalman filtering system to evaluate the impact of these measurements on analyses and forecasts of tropical storms. The RO refractivity observations are assimilated with other (routine) satellite data using a 36 km resolution WRF model. Results are presented for hurricane Dean of 2007 and typhoon Shanshan of 2006. The RO observations reduce the bias of the water vapor analysis and improve the ensemble analyses and forecasts of the intensity and position of the tropical storms.

347 Early Treatment Discontinuation: A Statistical Investigation ●▲

Biopharmaceutical Section, WNAR, Biometrics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Joint Modeling of a Longitudinal Outcome and Early Study Discontinuation

✱Donald Hedeker, University of Illinois at Chicago, School of Public Health (MC 923), 1603 W. Taylor St., room 955, Chicago, IL 60612-4310, *hedeker@uic.edu*

Key Words: joint model, mixed model, longitudinal

Joint models for longitudinal data and survival outcomes have been increasingly used, in particular in longitudinal studies where dropout occurs. Here, we will present an example of such modeling for a longitudinal psychiatric study comparing drug to placebo, in which a fair proportion of the subjects did not complete the study. Augmenting a standard mixed model for longitudinal data, we will specify a grouped-time survival model for study discontinuation that depends on the random effects of the longitudinal model. This type of model has been described as a shared parameter model by Ten Have (1996). In particular, we will illustrate the importance of including interactions of the random effects with treatment group in the model for study discontinuation to account for potentially different reasons for dropout. Application is described for several types of longitudinal outcomes.

Predictors of Premature Treatment Discontinuation: A Statistical Investigation Across Four Disease States

✱Hong Liu-Seifert, Eli Lilly and Company, *Liu-Seifert_hong@lilly.com*

Key Words: discontinuation, attrition, Cox Regression, Time dependent covariate, predictor

The traditional perception has been that early treatment discontinuation is driven mainly by unwanted side effects. This research aims to challenge this wisdom by utilizing statistical tools. Cox regression models with time dependent covariates were used to assess the predictive values of multiple independent variables including treatment response and side effects across different disease states. Inadequate early treatment response measured at first post-baseline visit was a significant predictor of patient discontinuation for all four disease states. At any visit during the studies, inadequate efficacy was the primary risk factor for early discontinuation with medication intolerability as the secondary driver. Treatment response is a more influential predictor of early attrition than adverse events across the four disease states studies.

Reducing Attrition Bias with an Adjustment for Participants' Intent To Dropout of a Clinical Trial

✱Andrew C. Leon, Weill Medical College of Cornell University, *acleon@med.cornell.edu*; Hakan Demirtas, University of Illinois at Chicago; Donald Hedeker, University of Illinois at Chicago

Key Words: attrition, clinical trial, mixed-effects model, dropout, Intent to Attend

Although mixed-effects models can accommodate incomplete data, the assumption of ignorable attrition is usually required for valid inferences. In an effort to fulfill the ignorability assumption, we consider a covariate that has been recommended by others, asking participants to rate their Intent to Attend the next assessment session. A simulation study compares the bias and coverage in mixed-effects outcome analyses that do and do not include Intent to Attend as a covariate. For the simulation specifications examined, if time-varying Intent to Attend variable is associated with attrition, outcome and treatment group, bias is substantially reduced by including it the outcome analyses. Accounting for participants' Intent to Attend the next assessment session can reduce attrition bias under conditions examined here.

348 Analysis of Mixed Outcome Data in Health and Medicine via Copula Models ●

Biometrics Section, Biopharmaceutical Section, Section on Statistics in Epidemiology, SSC, WNAR, Section on Health Policy Statistics

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Do Prescription Drug Expenditures Have Cost-Offsets? A Panel Data Analysis Using Copulas

✱Pravin K. Trivedi, Indiana University, 105 Wylie Hall, Dept of Economics, 100 S. Woodlawn Av, Bloomington, IN 47405-7104, *trivedi@indiana.edu*; David Zimmer, Western Kentucky University

Key Words: Panel data, Likelihood inference, cost offsets

There is a body of literature that considers the question whether policymakers should be concerned about the rapid growth of prescription drug expenditures or whether such expenditures on prescription drugs pay for themselves by offsetting other healthcare expenditures. The empirical results at present are inconclusive. Using the framework of bivariate copulas, the paper considers whether prescription drugs and other medical care services are substitutes or complements in consumption. Two nationally representative panel datasets are used: the Medical Expenditure Panel Survey (MEPS) and the Medicare Current Beneficiaries Survey (MCBS).

Modeling Count Data with Copulas: Shall We?

✱Johanna Neslehova, ETH Zurich, Departement Mathematik, Raemistrasse 101, Zurich, International 8092 Switzerland, *johanna@math.ethz.ch*; Christian Genest, Université Laval

Key Words: copula, discrete distribution, kendall's tau, spearman's rho, count data

Various facts will be reviewed about copulas linking discrete distributions. It will be shown that the possibility of ties resulting from atoms in the probability distribution invalidates various familiar relations that lie at the root of copula theory in the continuous case. The talk will highlight some of the dangers and limitations of undiscriminating transposition of modeling and inference practices from the continuous setting to the discrete one.

Efficient Bayesian Inference for Gaussian Copula Regression Models

✱Michael K. Pitt, University of Warwick, CV47AL, Coventry, International CV47AL United Kingdom, *m.pitt@warwick.ac.uk*

Key Words: copula, markov chain monte carlo

A Gaussian copula regression model gives a tractable way of handling a multivariate regression when some of the marginal distributions are non-Gaussian. Our article deals with Bayesian inference for a Gaussian copula model and makes the following contributions. First, it presents a general Bayesian approach for estimating a Gaussian copula model that can handle any combination of discrete and continuous marginals. Second, it generalizes Gaussian graphical models to the Gaussian copula framework by using a prior that allows the off- diagonal elements of the inverse of the correlation matrix to be identically zero. Third, the article constructs a novel and efficient simulation method for carrying out posterior inference. Fourth, the methods in the paper are applied to simulated and real data sets and are demonstrated to work well empirically.

349 Statistical Methods for Differential Equation Models ●

Section on Physical and Engineering Sciences, IMS
Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Differential Equation Modeling of HIV Viral Fitness Experiments: Model Identification, Model Selection, and Multimodel Inference

✱Hongyu Miao, University of Rochester, Department of Biostatistics and Computational Biology, 601 Elmwood Ave, Box 630, Rochester, NY 14642, *Hongyu_Miao@urmc.rochester.edu*; Hulin Wu, University of Rochester; Carrie Dykes, University of Rochester; Lisa M. Demeter, University of Rochester

Key Words: Differential Equation Modeling, HIV Viral Fitness, Global Optimization, Identifiability Analysis, Model Selection, Multi-model Inference

Many biological processes and systems can be described by a set of differential equation models. However, literature in statistical inference for differential equation models is very sparse. We propose statistical estimation, model selection and multi-model averaging methods for HIV viral fitness experiments in vitro that can be described by a set of nonlinear ordinary differential equations (ODE). The parameter identifiability of the ODE models is also addressed. We apply the proposed methods and techniques to experimental data of viral fitness for HIV-1 mutant 103N. We expect that the proposed modeling and inference approaches for the differential equation models can be widely used for a variety of biomedical studies.

Inverse Modeling of Full-Waveform, Single-Well Geophysical Data Using a Bayesian Model and Markov Chain Monte Carlo Methods

✱Jinsong Chen, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS90-1116, Berkeley, CA 94720, *jchen@lbl.gov*; Thomas M. Daley, Lawrence Berkeley National Laboratory; Carlos Torres-Verdin, University of Texas at Austin

Key Words: geophysical inversion, elastic wave equations, Bayesian, MCMC

Single-well seismic methods have been used for environmental investigations and natural resources exploration as cost-effective approaches for subsurface imaging. Current modeling studies primarily focus on numerical simulation of the governing elastic wave equations. In this study, we develop a Bayesian inversion method to estimate elastic parameters and their spatial distribution in the surrounding medium from full-waveform seismic data. We use a staggered-grid finite difference method to forward model full seismic waveforms in a 2D cylindrical coordinate system. We use Metropolis-Hastings and slice sampling methods to explore the joint posterior probability distribution function. Both synthetic and field case studies show that the developed Bayesian model is effective for estimating elastic properties as well as their associated uncertainty.

Data Mining Dynamical Systems: Automated Symbolic System Identification for Exploratory Analysis

✱Michael D. Schmidt, Cornell University, 32 Fairview Sq, Ithaca, NY 14850, *mds47@cornell.edu*; Hod Lipson, Cornell University

Key Words: Dynamical Systems, Symbolic Regression, Genetic Programming

Many branches of science and engineering represent dynamical systems mathematically, as sets of differential equations, derived laboriously from basic principles and experimentation. We are developing new approaches to reverse-engineer the analytical differential equations of a dynamical system automatically. We use genetic programming techniques to perturb and destabilize the system to reveal its hidden characteristics and to infer nonlinear symbolic relationships between variables. Our research has shown the ability to infer a seven-variable cell glycolysis system directly from data - the largest system inferred automatically to date. Our focus is to advance this promising approach to operate in high-noise and limited observability environments where manual methods for modeling are most overwhelmed.

An Approximate Maximum Likelihood Method for Parameter and State Estimation in Continuous-Time Dynamic Models

✱P. James McLellan, Queen's University, Department of Chemical Engineering, Kingston, ON K7L 3N6 Canada, *james.mclellan@chee.queensu.ca*; Saeed Varziri, Queen's University; Kim B. McAuley, Queen's University

Key Words: differential equation models, parameter estimation, maximum likelihood, state estimation, noise variance estimation, spline regression

Parameter estimation for continuous-time dynamic models of chemical processes must deal with unknown initial conditions, poorly parameterized models, measurement and state disturbances, irregularly sampled observations, parameter nonlinearity and partial state measurements. An approximate maximum likelihood (AMLE) formulation has been developed to address these challenges that approximates the solution of the nonlinear stochastic differential equation models using suitable basis functions (e.g., splines). Measurement noise variance is often known, but disturbance variances are typically unknown. A more stable method for estimating disturbance variance is proposed, and the AMLE technique is illustrated by estimating parameters in two dynamic models of chemical reactor processes—a nylon reactor model and a stirred tank chemical reactor simulation.

350 Multimarker Analysis in Genetic Association Studies ●

Section on Statistics in Epidemiology, Biometrics Editorial Board, WNAR, Biometrics Section
Wednesday, August 6, 8:30 a.m.–10:20 a.m.

A Constrained Regression Approach for Studying Haplotype-Specific Effects

✱Jung-Ying Tzeng, North Carolina State University, Campus Box 7566, Dept of Statistics, Raleigh, NC 27513, *jytzeng@stat.ncsu.edu*; Howard Bondell, North Carolina State University

Key Words: haplotype analysis, constrained regression, genetic association study

Understanding the effects of specific haplotypes can help to identify etiological variants and infer biological explanations. Ideally, a thorough haplotype-specific analysis should call for pairwise comparisons among all haplotypes, similar to the post-hoc analysis of ANOVA. However, such comparisons may suffer from power loss due to multiple testing adjustment, and often yields contradictory conclusions on which haplotypes share the same level of effects. To overcome this concern, we propose a constrained regression approach that performs the ANOVA-type of post-hoc analysis for haplotype-specific analysis. The method uses constraints that encourage haplotypes with similar effect sizes to be estimated with exact equality, and transfers the posthoc analysis from a multiple-comparison procedure to a variable-selection framework.

Pathway-Based Analysis of Genomewide Association Studies

✱Kui Zhang, The University of Alabama at Birmingham, 1665 University Blvd., Dept. of Biostatistics, Ryals School of Public Health Bldg. 327H, Birmingham, AL 35294, *kzhang@ms.soph.uab.edu*; Douglas K. Childers, The University of Alabama at Birmingham

Key Words: Genomewide Association (GWA) Study, Pathway, Interaction, SNP, Linkage Disequilibrium (LD), Association

Typical analysis methods for GWA studies are based on single markers and generally ignore the rest of information. Wang et al. (AJHG, 81, 1279-1283) proposed a pathway-based analysis and illustrated that joint analysis of multiple genes in the same pathway can provide additional insights into interpretation of GWA studies. Although multiple SNPs may be genotyped for a gene, Wang's method needs to find a single SNP to represent all SNPs within a gene and the optimal choice of such a SNP is not clear. Wang's methods can not detect the joint effects of multiple SNPs within and between genes. We develop a pathway based method that can analyze all SNPs within a gene and joint effects of multiple SNPs. We apply our method to a published GWA study and demonstrate that the new developed method can provide additional findings in GWA studies.

A Retrospective Method for Inference on Haplotype Main Effects and Haplotype-Environment Interactions Using Clustered Haplotypes

✱Martha Jones, PRA International, 1602B Keith Valley Rd, Charlottesville, VA 22903, *marti.jones@gmail.com*; Michael Epstein, Emory University; Glenn Satten, Centers for Disease Control and Prevention; Andrew Allen, Duke University Medical Center; Jung-Ying Tzeng, North Carolina State University

Key Words: haplotype analysis, retrospective likelihood, haplotype clustering, multi-marker analysis, case-only analysis, evolutionary clustering

Regression-based methods for haplotype association analysis in case-control studies are generally based on either a prospective or retrospective framework. For haplotype analysis, previous work has shown that a retrospective framework can be more efficient than a prospective one. This work extends the concept of haplotype clustering to the retrospective framework, and improves the performance of haplotype analysis on main and HxE effects in case-control studies. Specifically, we construct a retrospective likelihood that allows for environmental covariates and HxE interactions. We derive generalized score statistics to test for main and interaction effects at the global and individual levels. Through simulation, we assess the validity of the proposed tests and, where appropriate, compare the power with that of the retrospective full-dimensional and prospective analyses.

Haplotype Associations When Functional Variation Is Only Partially Scored

✱Dmitri Zaykin, NIEHS / NIH, Building 101 / Room B356B, P.O. Box 12233; Mail Drop A3-03, Research Triangle Prk, NC 27709, *zaykind@niehs.nih.gov*

Key Words: haplotype association, linkage disequilibrium

We consider an association mapping scenario where scored haplotypes are either linked to causal variants via linkage disequilibrium (LD), or represent only a subset of all causal polymorphisms. In both cases, trait variances among the observed haplotypes may differ from each other. Such haplotypic variance contrast can be estimated, and incorporated into haplotypic tests for association for unphased data. The inference involves simultaneous estimation of haplotypic effects, their frequencies, and haplotype-specific variances. The approach provides additional power under models where only a subset of functional polymorphisms has been scored, as well as under heterogeneity models, where multiple unobserved mutations are linked to non-functional observed polymorphisms via LD. A pronounced variance contrast may indicate such involvement of additional factors that influence the trait.

A New Support Vector Regression Approach to Gene Selection

✱Pei-Chun Chen, National Taiwan University, No. 17 Xu-Zhou Road, Taipei 100, Taiwan, Taipei, 100 R.O.C., *d93842005@ntu.edu.tw*; Su-Yun Huang, Institute of Statistical Science, Acadamia Sinica; Chuhsing K. Hsiao, National Taiwan University

Key Words: SNP, support vector machine, support vector regression

Genetic data are often collected with a large number of genes and a relatively small simple size. However, only a few genes among them are associated with the disease of interest. As the number of markers increases, the statistical analysis becomes complicate due to the dimensionality. A great challenge is to find the genes showing significant association with the disease. Most approaches focus on a multi-hypotheses testing of the gene-disease association. Recently, the statistical machine learning methods are applied to this problem, such as support vector machine (SVM). We propose a new approach for selection by weighting on every sample differently. These weights are calculated via support vector regression. A SNPs data set of schizophrenia will be analyzed for illustration.

351 Statistical Literacy 2008▲

Section on Statistical Education

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Interpreting the Substantive Significance of Multivariate Regression Coefficients

✱Jane E. Miller, Rutgers, The State University of New Jersey, Institute for Health Research, 30 College Avenue, New Brunswick, NJ 08901, *jmiller@ifh.rutgers.edu*

Key Words: multivariate regression, statistical significance, substantive significance, writing

A critical objective for applications of multivariate regression analysis is evaluation of both substantive importance and statistical significance, yet many articles focus excessively on inferential statistical tests at the expense of substantive issues. In this presentation, I demonstrate approaches for writing clear sentences to interpret the substantive meaning of coefficients from ordinary least squares regression, taking into account the type of independent variable and the distributions of the dependent and independent variables. After introducing "the Goldilocks principle"—no one size contrast fits all variables—I use diverse examples to illustrate the importance of considering both the topic and data when evaluating substantive significance. Complementary use of prose, tables and charts to present both statistical and substantive significance will also be covered.

Just Plain Data Analysis

✱Gary Klass, Illinois State University, 4600 Politics and Government, Illinois State University, Normal, IL 61790-4600, *gmklass@ilstu.edu*

In the discipline of political science, and other social science and professional disciplines as well, the ongoing debate between adherents to quantitative and qualitative methodologies ignores the most prevalent form of social science quantitative methodology. Just plain data analysis involves essential statistical literacy skills of finding, presenting and interpreting social science data that are common to almost all of the discipline's subfields and that are not often taught in social science research methods and statistics courses.

Using Simulation To Teach Statistical Literacy

✱Marc Isaacson, Augsburg College, 1964 Juliet Ave, Saint Paul, MN 55105, *isaacson@augsburg.edu*

One aspect of statistical literacy is the ability to read and interpret survey data presented in tabular and graphical formats. In addition, some may include the tasks of data collection and the creation of such tables and graphs under the auspices of statistical literacy. Unfortunately, properly collecting survey data may take a great deal of time. This paper presents a unique simulation program that allows students to select their topic of interest and the associated survey questions. Given student input, the simulation program generates a random sample of simulated survey data which the student is expected to analyze and communicate in ordinary English. In addition, the program is flexible enough to allow students to work with the generated raw data to become active participants in the creation of tables and graphs. Preliminary results are presented along with student feedback.

Numbers in the News: A Survey

Milo Schield, W.M. Keck Statistical Literacy Project; ✱Robert L. Raymond, University of St Thomas, 800 Grandview Ave W, Roseville, MN 55113, *rlraymond@stthomas.edu*

Key Words: statistical literacy, statistical education

Many teachers of introductory statistics state that producing statistically literate students is one of their goals. When asked to elaborate, they may say that they want their students to read the newspaper critically or with deeper understanding if they encounter articles with statistical content. It would seem, then, that those teachers would find it useful to know the frequency with which various statistical concepts or skills are needed to read newspaper reports containing numerical results. This paper surveys over 500 news articles, classifying and tabulating the statistical knowledge into more than 100 categories that are relevant to understanding them.

Von Mises' Frequentist Approach to Probability

✱Milo Schield, W.M. Keck Statistical Literacy Project, 2211 Riverside Avenue, Augsburg College, Minneaoplis, MN 55454-1350, *milo@pro-ns.net*; Tom Burnham, Cognitive Consulting

Key Words: statistical literacy, statistical education

Richard von Mises formulated a strict Frequentist approach to probability. This approach is limited to open-ended data sets for which there are sufficient reasons to project future stability—to believe that the relative frequency of the observed attribute would tend to a fixed limit if the observations were continued indefinitely. Unlike conventional mathematical theories which are of abstract origins, von Mises' approach is intended to make predictions which are true of the real world. For von Mises, probability statements are never applicable to non-random processes, to unstable processes, to a single individual, or to outcomes that are untestable (e.g., Bayesian strength of belief). Many—if not most—statements involving chance, odds, risk, likelihood and probability along with comparisons involving likely, risky, and probable are improper under this strict Frequentist approach.

352 Design Issues for Conjoint Experiments ●

Section on Statistics and Marketing

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

A Comparison of Bayesian Design Criteria

✱Jie Yu, Katholieke Universiteit Leuven, Naamsestraat 69, Leuven, International B-3000 Belgium, *Jie.Yu@econ.kuleuven.be*; Peter Goos, Universiteit Antwerpen; Zsolt Sándor, Universidad Carlos III Madrid; Martina Vandebroek, Katholieke Universiteit Leuven

Key Words: Bayesian design criterion, posterior density, expected posterior covariance, conjoint choice design, Laplace approximation, Fisher information

The Bayesian design criterion often used in the literature is derived from the Hessian matrix of the log likelihood function. However, other design criteria are possible. Examples are design criteria based on the Hessian matrix of the log posterior density, on the expected posterior covariance, or on the amount of information provided by the experiment. In this study, we apply these criteria to choice experimental designs. We investigate how robust the resulting Bayesian optimal designs are with respect to other design criteria for which they are not optimized and examine which design criterion is most appealing in a non-Bayesian framework where it is accepted that prior information must be used for design but should not be used in the analysis, and which one is most appealing in a Bayesian framework when the prior distribution is taken into account for design and for analysis.

Adaptive Self-Explication of Multi-Attribute Preferences

✱Oded Netzer, Columbia University, Columbia Business School - Uris Hall 520, 3022 Broadway, New York, NY 10027, *on2110@columbia.edu*; V. Seenu Srinivasan, Stanford University

Key Words: Preference Measurement, Conjoint Analysis, Self-explicated, Survey Design, Constant Sum Questions, Adaptive Questionnaires

We propose a computer-based adaptive self-explicated approach for multi-attribute preference measurement. This approach overcomes some of the limitations of the ratings and constant sum allocation approaches commonly used to estimate the attribute importance in self-explicated applications. We developed a computer-based approach that breaks down the attribute importance question into a rank-order task plus a sequence of constant sum, paired comparison questions. The sequence of questions is chosen adaptively for each respondent to maximize the information elicited from each paired comparison. We find that the predictive validity of the proposed approach is 38%-52% higher than that of the Adaptive Conjoint Analysis, the Fast Polyhedral approach, and the traditional self-explicated method. Additionally, the proposed approach allows for mitigation of the respondents' burden.

Designing Optimal Discrete Choice Experiments for Alternative-Specific Attributes

✱Leonie Burgess, University of Technology, Sydney, PO Box 123, Broadway NSW, International 2007 Australia, *leonie.burgess@uts.edu.au*

Key Words: choice experiments, alternative-specific, optimal designs

Discrete choice experiments (DCEs) have been used in marketing, transport and health economics, among other areas, for more than 20 years. In DCEs the attributes can either be generic (unlabelled) or they can be alternative-specific (labeled). There is a substantial body of literature on designing optimal or, at least, efficient DCEs for generic attributes, but there has been little to date on the optimality of DCEs with alternative-specific attributes. Traditionally, DCEs with alternative-specific attributes have been constructed using one fractional factorial design with M x A attributes in total, where A attributes describe each of M options. However, until now, nothing has been known about the optimality properties of these designs, when used to estimate alternative-specific effects. In this paper we present optimality properties of DCEs with alternative-specific attributes.

Efficient Experimental Designs for Hyperparameter Estimation: Studying the Level-Effect in Conjoint Analysis

✱Qing Liu, University of Wisconsin-Madison, Madison, WI 53706, *qliu@bus.wisc.edu*; Angela M. Dean, The Ohio State University; David Bakken, Harris Interactive; Greg Allenby, Then Ohio State University

Key Words: Design efficiency, Conjoint Analysis, Hyperparameters

Research in marketing, and business in general, involves understanding when effect-sizes are expected to be large and when they are expected to be small. An example is the level-effect in marketing, where the effect of product attributes on utility is positively related to the number of levels present among choice alternatives. Knowing the contexts in which consumers are sensitive to the levels of attributes is an important aspect of merchandising, selling and promotion. In this paper, we propose efficient methods of learning about contextual factors that influence consumer preference and sensitivities within the context of a hierarchical Bayes model. A design criterion is developed for hierarchical linear models, and validated in a study of the "level-effect" in conjoint analysis using a national sample of respondents. Extensions to other model structures are discussed.

Optimality Criteria for Designs To Measure Accurately the WTP by Conjoint Choice Experiments

✱Bart Vermeulen, K.U.Leuven, Naamsestraat 69, Leuven, Belgium, *bart.vermeulen@econ.kuleuven.be*; Peter Goos, Universiteit Antwerpen; Martina Vandebroek, Katholieke Universiteit Leuven

Key Words: experimental design, willingness-to-pay, conjoint choice experiments

In the paper "How to measure efficiently the willingness-to-pay with conjoint choice experiments," the authors Vermeulen et al. present an optimality criterion to generate designs for conjoint choice experiments to measure accurately the willingness-to-pay (WTP). This criterion was based on minimizing the variance of the WTP estimator. The Delta method they used to approximate this variance is not very accurate. Furthermore, a poorly estimated price coefficient often leads to unrealistic estimates of the WTP. In this paper, other optimality criteria are considered to solve those two shortcomings. These criteria are based on different approaches to estimate confidence intervals for the WTP. We compare the accuracy of the WTP estimators resulting from the designs generated by these new criteria with the results of the previous paper.

353 Recent Developments in Transportation Statistics ●▲

Social Statistics Section, Section on Government Statistics, Section on Survey Research Methods
Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Variations in Effect Estimation Techniques To Assess Graduated Driver

✱David M. Neyens, The University of Iowa, 3131 Seamans Center, Iowa City, IA 52242, *dneyens@engineering.uiowa.edu*; Birsen Donmez, Massachusetts Institute of Technology; Linda Boyle, The University of Iowa

Key Words: teen drivers, crash data, time-series models, graduated licensing programs

The goal of Graduated Drivers Licensing programs is to provide teenage drivers the ability to gain driving experience while limiting their exposure to the most dangerous driving situations. The examination of these programs has been conducted using crash data and with various statistical techniques leading to very different estimates for GDL effectiveness. Failure to account for time and seasonal effects can result in overestimation of this public health intervention. This study uses data from Iowa as an example to demonstrate how various methods can have very different results. Time series models tend to provide the best estimate when compared to general linear models, relative risks, and rate ratios.

Quasi-Likelihood Generalized Linear Regression Analysis of Fatality Risk Data

✱C. Craig Morris, Bureau of Transportation Statistics, 1200 New Jersey Ave, SE/RITA/USDOT, Washington, DC 20590, *craig.morris@dot.gov*

Key Words: quasi-likelihood, regression, Poisson, overdispersion, motorcycle, helmet

Transportation-related fatality risk is a function of many interacting human, vehicle, and environmental factors. Statistically valid analysis of such data is challenged both by the complexity of plausible structural models relating fatality rates to explanatory variables and by uncertainty about the probability distribution of the data. But, fortunately, generalized linear modeling and maximum quasi-likelihood estimation together provide an extraordinarily effective set of statistical tools for the analysis of such data. The goal of this talk is to illustrate and promote application of these tools to fatality risk analysis.

Estimating Nonresponse Bias in the Omnibus Household Survey

✱Promod Chandhok, RITA/BTS, U.S. Department of Transportation, 1200 New Jersey Avenue, SE, Washington, DC 20590 USA, *promod.chandhok@dot.gov*

Key Words: call history, travel surveys, response error

The Omnibus survey is a random digit dial survey of slightly more than 1000 noninstitutionalized telephone households in U.S. Information is collected to monitor expectations of and satisfaction with the transportation system. Since the response rate for the survey is approximately 50%, there has been considerable interest in gauging the nonresponse bias in this survey. There are mainly three options for assessing nonresponse bias. The first option, used frequently, is an intensive follow up of nonrespondents. This was not an option here because of cost. The second option is external validation—benchmarking against some other data, and this was done here for commonly used demographic variables. This operation reduced bias. Another option is to use data on survey operations. This presentation discusses results from using call history data to estimate nonresponse bias in key variables.

Safety Analyses of Signalized Intersections Using Bayesian Hierarchical Spatial Models

✱Feng Guo, Virginia Polytechnic Institute and State University, 415 B Hutcheson Hall, Dept. of Statistics, Blacksburg, VA 24061, *feng.guo@vt.edu*; Xuesong Wang, University of Central Florida; Mohamed Abdel-Aty, University of Centraol Florida

Key Words: Bayesian, Spatial, Signalized intersection, CAR, Traffic Safety

Signalized intersections are among the most dangerous locations on a highway network. There are more than 2.8 million intersection-related crashes in year 2000. In this paper we propose a Bayesian hierarchical approach to model accidents at signalized intersections. The data include 197 signalized intersections in the State of Florida. The proposed model incorporates the effects of geometric design features, traffic control features, and traffic characteristics. Furthermore, we argue that signalized intersections along a certain corridor will affect each other, especially for those in close proximity. We incorporate this spatial correlation into the model by using a conditional autoregressive prior for an intersection specific random effect. The results of this study will benefit intersection geometric design and traffic management.

Analyzing Time Delay of Highway Construction Transportation Contracts: An Exploratory Study

✱Abhishek Bhargava, Purdue University, 625 N 4th St., Apt 2, lafayette, IN 47901, *abhargava@purdue.edu*; Panagiotis Anastasopoulos, Purdue University; Fred Mannering, Purdue University; Kumares C. Sinha, Purdue University; Samuel Labi, Purdue University

Key Words: Time Delay, Construction contracts, Hazard Models, Logit Models

Many highway agencies face the problem of contract time delays in transportation construction projects. Such delays can have adverse effects on the owner, contractor and the facility users. Time delay may be caused by any party to the contract; its causes may include changes in engineering design and the factors of production, and certain characteristics of the contract bidding process and project environment. Focusing on the latter, in this study a methodology is developed to investigate the likelihood and extent of time delay in highway construction projects. Logit and hazard-based duration models are estimated using data from a state agency, and including explanatory factors relating to project type, project environment and contract bidding characteristics. The demonstrated methodology can help agencies identify the factors that contribute to time delays and to better predict them.

354 Variance Estimation ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Single Phase Variance Estimation Approach to Two-Phase Designs

✱Avinash C. Singh, Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *avi.singh@statcan.ca*

Key Words: Two phase Variance estimation, Across PSU Conditionally Independent Designs, With Replacement PSU Assumption

Two phase designs are useful for increasing estimation efficiency for reasons including deep stratification, improved nonresponse adjustment and calibration controls, and for integrating survey designs by using the first phase as a master sample. Despite these benefits, such designs are not in wide use due to difficulty in variance estimation unlike the multi-stage single phase designs where the WRPSU assumption is often invoked for simplified variance estimation. It is shown that by a simple modification of two phase cluster designs in which the second phase sampling within each first phase cluster/PSU is performed pps independently across PSUs conditional on the first phase sample, the usual simplified variance estimation remains applicable even if the second phase variance is not estimable. The size measures for the pps correspond to sampling rates determined at the second phase.

Some Notes on Cell Collapsing and Its Effect on Replicate Variance Estimates with the Delete-a-Group Jackknife Variance Estimator

✱Katherine J. Thompson, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, *katherine.j.thompson@census.gov*; Samson Adeshiyan, U.S. Census Bureau

Key Words: cell collapsing, quasi-randomization, variance estimation

Often surveys employ weight adjustment procedures to compensate for unit non-response. Cell collapsing is typically performed when the respondent counts within individual unit nonresponse adjustment cells are small. Additionally, surveys may impose a collapsing criteria based on the magnitude of the adjustment factor. For replicate-based variance computation, there are few guidelines for which weighting cells should be used for the replication process (i.e., whether to use the weighting cells from the full sample in all replicates or to determine weighting cells individually in each replicate sample). We address how weighting cell collapsing should be handled for the simple delete-a-group jackknife replicate variance estimation procedure when using the "quasi-randomization" estimator to account for unit nonresponse via a simulation study modeled after a typical business survey.

Variance Estimation for an Estimator of Between-Year Change in Totals from Two Stratified Bernoulli Samples

✱Kimberly A. Henry, Internal Revenue Service, P.O. Box 2608, Washington, DC 20013-2608, *Kimberly.A.Henry@irs.gov*; Valerie L. Testa, Internal Revenue Service; Richard Valliant, The University of Michigan

Key Words: Horvitz-Thompson estimator, post-stratification estimator, survey sampling

This paper provides the theoretical framework for estimating the variance of the difference in two years' totals estimated under the stratified Bernoulli sample design. We provide a design-unbiased estimator that accounts for two practical problems: a large overlap of sample units between two years' samples and "stratum jumpers," which are population and sample units that shift across strata from one year to another. Both problems affect estimating the covariance term in the variance of the difference. The estimator is applied to data from the Statistics of Income Division's individual tax return sample. Naïve variance estimates using only the separate years' variances are compared to show the effect of ignoring the estimated covariance.

Variances for the American Community Survey 2005 Public Use Microdata Sample

✱Sirius C. Fuller, U.S. Census Bureau, 3927 Madison St, Hyattsville, MD 20781-1750, *sirius.c.fuller@census.gov*; Anthony G. Tersine, Jr., U.S. Census Bureau

Key Words: American Community Survey, Generalized Variance Function, Public Use Microdata Sample, Replicate Weights

The American Community Survey (ACS) releases data every year including a Public Use Microdata Sample (PUMS) file. PUMS allows data users to tabulate their own characteristic estimates by publishing responses from individual households and persons after removing all identifying information to ensure confidentiality. There are two methods available for users to calculate standard errors: generalized variances (design factors) and replicate weights (direct standard errors). Standard errors calculated using replicate weights are more accurate than those using design factors. Previous research determined that design factors are an acceptable approximation for direct standard errors for totals. Using that as a benchmark, this paper investigates whether the design factors for means and medians are an acceptable substitute for their direct standard errors and proposes alternative options.

Estimating Observation Outlyingness Using Statistical Distances in National Resources Inventory

✱Jianqiang Wang, Iowa State University, Center for Survey Statistics and Methodology, 2625 N. Loop Dr., Bldg. 2, Suite 2140, Ames, IA 50011, *jqwang@iastate.edu*; Jean Opsomer, Colorado State University

Key Words: outlier identification, longitudinal survey, statistical distance

A common issue in large-scale complex surveys is the detection of outliers in the data. Such outliers can be caused by frame imperfections or by errors during data collection. The existence of the outliers in a survey dataset can cause bias in the inference for population quantities of interest. We describe a procedure to identify suspicious points in survey data, taking into account the structure of the finite population. We compute the distance from each point in the dataset to the center of a subpopulation that it belongs to, and define a measure of point outlyingness as the tail probability on the resulting distance distribution function. We also extend the procedure to periodically updated surveys when detection rules are built based on historical data and applied to recently collected data. We apply the proposed approach to National Resources Inventory, a large-scale longitudinal complex survey.

An Alternative to the Logit-Wald Method for Inference with Models for Proportions

Avinash C. Singh, Statistics Canada; ✱Claude Nadeau, Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6, *claude.nadeau@statcan.ca*

Key Words: Estimating functions, confidence intervals, variance estimation, survey data

The quasi-likelihood estimating function (EF) properly weighted for survey data can be used to fit logit models. In practice, an asymmetric logit-Wald confidence interval (CI) is built due to skewness of estimates of low or high prevalence outcomes. However this method has instability in the variance estimate for moderate sample sizes and/or for low or high prevalence outcomes, leading to poor coverage properties. We propose an alternative based on the randomly recentered estimating equations (RREE) idea of Singh and Dochitoiu (2005). Replicate parameter estimates are created by equating the standardized EF to random values drawn from the N(0,I) distribution. They are used to compute new variance and interval estimates. Simulations are used to evaluate the performance of RREE relative to logit-Wald in terms of bias and variance of the variance estimator, length and coverage of the CI.

355 Weighting in Longitudinal and Multilevel Surveys ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Creating Bootstrap Survey Weights When Only Limited Design Information Is Available: A Simulation Study

✱Michael Wendt, Statistics Canada, Social and Aboriginal Statistics Division, 7th Floor, Jean Talon Building, 170 Parkdale Avenue, Ottawa, ON K0A 1T6 Canada, *wendmic@statcan.ca*; David Paton, Statistics Canada

Key Words: survey weights, bootstrap weights, estimation in complex surveys

Statistics Canada's General Social Survey is an annual cross-sectional household survey for which 20 complete cycles of data now exist. Recent cycles provide a final calibrated survey weight as well as sets of bootstrap weights for variance estimation. Analysts have requested bootstrap weights for earlier cycles, but after extensive research, we have determined that it is not feasible to construct them for the earlier cycles using current methods. However, detailed design information and final weights for each cycle do exist. We present a proposed method of creating "approximate" bootstrap weights for past cycles and evaluate them with a simulation study. Although this describes a solution to a specific problem, our method may be of wider use: the possibility of creating sets of pseudo-design based final bootstrap survey weights when only some limited information is available.

Minimizing Conditional Global MSE for Health Estimates from the Behavioral Risk Factor Surveillance System for U.S. Counties Contiguous to the United States-Mexico Border

✱Joe Fred Gonzalez, Jr., Centers for Disease Control and Prevention, NCHS, ORM, 3311 Toledo Rd., Room 3121, Hyattsville, MD 20782, *Jgonzalez@cdc.gov*; Machell Town, Centers for Disease Control and Prevention; Jay J. Kim, National Center for Health Statistics

Key Words: mean square error (MSE), unbiased estimation, poststratification, collapsing

The Behavioral Risk Factor Surveillance System (BRFSS) is a state-based telephone survey of the adult civilian non-institutionalized population residing in the United States. Consequently, the BRFSS final weights that are currently available in the public use data files are designed to produce unbiased estimates of health conditions by socio-demographic characteristics at the State level. In addition to State level BRFSS estimates, there is interest in the health status of adults residing in the 25 U.S. counties contiguous to the United States - Mexico border. The purpose of this paper is to apply an alternative approach for post-stratification by minimizing conditional global mean square error of BRFSS health estimates for adults residing in the 25 counties contiguous to the United States - Mexico border.

Results from the 2006 Canadian Census Weighting Process

✱Wesley Benjamin, Statistics Canada, *wesley.benjamin@statcan.ca*; Darryl Janes, Statistics Canada

Key Words: census, household weights, sampling bias, discrepancies, calibration, regression estimation

In the Canadian Census, basic person and dwelling information is gathered on a 100% basis, and additional information is collected from a 20% sample of private households. Initial weights of approximately five are applied to the sampled population but are then calibrated independently for approximately 6,600 geographical regions through a multistep regression estimation process to produce the final weights. Within each region, auxiliary variables are discarded from the weighting process for being collinear, nearly collinear or causing outlier weights. The weighting process is evaluated by comparing the estimates based on the final weights with the 100% totals for all the auxiliary variables. In addition, biases in the census sample for the auxiliary variables are studied by comparing estimates based on the initial weights to the 100% totals at both the national and regional levels.

Changing the Weighting Approach of a Longitudinal Survey

✱Michael Yang, ICF International, 9300 Lee Highway, Fairfax, VA 22031, *myang@icfi.com*; Yongyi Wang, National Opinion Research Center

Key Words: weighting class adjustment, response propensity weighting, longitudinal survey

The Survey of Doctorate Recipients (SDR) is a longitudinal survey of individuals with doctoral training in science and engineering in the US. The SDR has used a weighting class adjustment approach to compensate for unknown eligibility and nonresponse. This approach incorporates only a small set of variables in defining the weighting classes. Changing to a response propensity based weighting approach would help overcome this limitation. Under the new approach, eligibility and nonresponse weighting adjustments would be based on propensity scores predicted by multivariate models. This study examines the pros and cons of such a change in weighting methodology. It examines (1) if additional bias reduction may be achieved under the new approach; (2) how the new approach would affect the variance of the weights; and (3)) how the new approach would affect the longitudinal data series.

Response Patterns Among New Businesses

✱Zhanyun Zhao, Mathematica Policy Research, Inc., 600 Alexander Park, Princeton, NJ 08540, *zzhao@mathematica-mpr.com*; Frank Potter, Mathematica Policy Research, Inc.; Yuhong Zheng, Mathematica Policy Research, Inc.

Key Words: Longitudinal survey, Stratified sample, Non-response adjustment, CHAID, Logistic regression

The Kauffman Firm Survey (KFS) is a national-level longitudinal panel survey of new businesses, gathering data on the resources and processes to initiate and to sustain businesses. The sample is a stratified random sample of firms starting in 2004, with over-sampling of high technology firms. The purpose of this paper is to identify the characteristics of respondents for the first and second follow up surveys (baseline data collection was completed in 2005–2006) and describe the non-response adjustment methods. Both cross sectional weights and longitudinal weights were developed to accommodate analytical objectives. We will also discuss the response patterns status for these two weights. The non-response adjustments are based on models using Chi-squared Automatic Interaction Detector (CHAID) and logistic regressions.

Multilevel Modeling in Large-Scale Assessments with Informative Sampling Weights

✱Frank Jenkins, Westat, 1650 Research Blvd., Rockville, MD 20850, *FrankJenkins@Westat.com*

Key Words: multilevel modeling, sampling weights, bias, power

Multilevel models have become common in large scale assessments with multistage sampling and unequal probabilities of selection, e.g. educational studies where longitudinal outcomes are observed for students nested within schools and districts. Often selection probabilities are related to outcomes resulting in informative weights correlated with outcomes and covariates. This can produce inefficient or biased estimates of fixed and random effects. Previous studies of sampling bias have examined simple 2-level models. We conducted a simulation study addressing issues of efficiency and bias for fixed effects and variances in 3-level models with predictors at all levels and varying levels of informativeness. Type I and Type II error were assessed for fixed effects. The goal is to provide guidance for large scale analyses when sampling weights may be much less than ideal.

Weighting Methods in School-Based Surveys

✱Ronaldo Iachan, Macro International, 11785 Beltsville Drive, Calverton, MD 2o705, *ronaldo.iachan@macrointernational.com*; Pedro J. Saavedra, Macro International; James Ross, Macro International

Key Words: Weighting, Post-stratification, School surveys

Data from school survey samples are typically weighted to account for unequal probabilities of selection, and non-response at the student and school levels. Weight adjustments are usually performed in two stages based on weighting class adjustments and on post-stratification. Weighting class adjustments are based on weighting cells defined by school or classroom using student enrollment data. Post-stratification adjustments capitalize on external data available for post-strata defined by grade, race/ethnicity and gender. This paper compares alternative approaches to weight adjustments using simulations and school survey data. In forming adjustment cells, we consider the role of potential interactions between class/school and demographic characteristics.

356 Testing and Estimation ●

Section on Statistics in Epidemiology, Biometrics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Three-Dimensional, Array-Based Group-Testing Algorithms

✱Hae-Young Kim, New England Research Institutes, 9 Galen Street, Watertown, MA 02472, *hkim@neriscience.com*; Michael Hudgens, The University of North Carolina at Chapel Hill

Key Words: Array, Group testing, Hierarchical, Sensitivity, Specificity, Predictive values

We derive the operating characteristics of three-dimensional array-based testing algorithms for case identification in the presence of testing error. The operating characteristics investigated include efficiency (i.e., expected number of tests per specimen) and error rates (e.g., sensitivity, specificity, positive and negative predictive values). The methods are illustrated by comparing the proposed algorithms with previously studied hierarchical and two-dimensional array algorithms for detecting recent HIV infections in North Carolina. Our results indicate that three-dimensional array-based algorithms can be more efficient and accurate than previously proposed algorithms in settings with test error and low prevalence.

A Hierarchical Latent Class Model for Evaluating Diagnostic Tests for Chlamydia Trachomatis in the Absence of a Gold Standard

✱Liangliang Wang, The University of British Columbia, 333-6356 Agricultural Road, Department of Statistics, Vancouver, BC V6T1Z2 Canada, *l.wang@stat.ubc.ca*; Nandini Dendukuri, McGill University; Alula Hadgu, Centers for Disease Control and Prevention

Key Words: Bayesian inference, conditional dependence, Chlamydia trachomatis, sensitivity, specificity

In the evaluation of diagnostic tests in the absence of gold-standard test, both traditional latent class models (TLCM) and their recent extensions, such as latent class models with random effects, assume that there is only one latent variable. Recently, we proposed hierarchical latent class models (HLCM) that assume that tests based on different biological phenomena measure different latent variables, which in turn measure the latent true disease status. Our main objective is to illustrate the use of HLCM using simulation and 7 diagnostic tests for detecting Chlamydia trachomatis. We show that HLCM outperform TLCM and the combined reference standard (CRS) by including multiple latent classes, random effects, and covariates. We also show how the CRS is problematic as it is based on a circular argument and is not uniquely determined when we have multiple imperfect tests.

Comparing Different Diagnostic Methods When Gold Standard Is Not Available

✱Nong Shang, Centers for Disease Control and Prevention, 1600 Clifton Rd., Atlanta, GA 30333, *nms6@cdc.gov*

Key Words: bivariate mxture model, latent variable, sensitivity, specificity

We often need to compare the sensitivities and specificities of two or more diagnostic methods in order to decide which method is the best choice under various circumstances. However, in many situations, there is no gold standard for us to correctly evaluate these quantities. One example is to compare the traditional TST (Tuberculin Skin Test) method with the newly developed QFT (QuantiFERON-TB) test in diagnosing latent tuberculosis infection (LTBI). We propose a bivariate mixture model approach using latent variables that contains information of the prevalence rates of LTBI to solve the problem. Under some general assumptions, the approach gives reasonable estimation of the corresponding sensitivities and specificities. It can even identify the situations that each of the methods may perform best. The method can also be used to estimate the underlying prevalence rate of the disease.

Transformation-Invariant and Nonparametric Monotone Smooth Estimation of ROC Curves

✱Pang Du, Virginia Polytechnic Institute and State University, 13000 Foxridge Ln, Apt J, Blacksburg, VA 24060, *pangdu@vt.edu*; Liansheng Tang, George Mason University

Key Words: ROC curve, transformation invariance, monotonicity, boundary constraints, nonparametric, confidence band

The accuracy of a diagnostic test in distinguishing diseased subjects from non-diseased ones is often evaluated by receiver operating characteristic (ROC) curves. Smooth ROC estimates are often preferable for continuous test results when the underlying ROC curves are continuous. Parametric methods need specific model assumptions. The existing nonparametric methods do not always satisfy the inherent properties of ROC curves, such as monotonicity and transformation invariance. In this article we propose a monotone spline approach to obtain smooth monotone ROC curves. Our method ensures important inherent properties of underlying ROC curves including monotonicity, transformation invariance, and boundary value constraints. We compared the newly proposed method with other ROC smoothing methods in large-scale simulation studies and illustrated our method through a real life example.

Assess the Diagnostic Accuracy of Sequential Screening Tests in a Multiphase Study of Dementia

✱Binbing Yu, National Institute on Aging, 7201 Wisconsin Avenue, Suite 3C309, Bethesda, MD 20904, *yubi@mail.nih.gov*; Chuan Zhou, Vanderbilt University

Key Words: Dementia, multi-phase design, ROC, screening, verification bias

In disease diagnosis when definite test is too invasive or expensive to be used to all subjects, a two-phase design is often used. The inexpensive Phase 1 test is used to determine the subjects that receive the gold standard. Analysis restricted only to verified cases lead to verification bias. Multiphase design is now often used in dementia and cancer diagnosis. The prevalent test in Phase 1 has high specificity, but relatively low sensitivity, Phase 2 consists of another prevalent test or a more confirmatory test, and the final test is the gold standard. Here, a method is proposed to estimate the ROC curves of each single test and the combined multiple tests in multi-phase design with verification bias. Performances of different schemes of combining repeated test are compared by simulation. The method is applied to data from a multiphase study of dementia.

Generalized ROC Criterion for Multivariate Normally Distributed Biomarkers with Limits of Detection

✱Neil J. Perkins, National Institute of Child Health and Human Development, 6100 Executive BLVD, Rockville, MD 20852, *perkinsn@mail.nih.gov*; Enrique F. Schisterman, National Institute of Child Health and Human Development; Albert Vexler, National Institute of Child Health and Human Development

Key Words: reciever operating characteristic curve, maximum likelihood, area under the curve, best linear combination, generalized ROC criterion, limit of detection

Authors have shown that a best linear combination (BLC) of multivariate normally distributed biomarkers can be formed parametrically that maximizes the area under the receiver operating characteristic curve (AUC). However, biomarkers are often measured with limits of detection (LOD) and the common practices, ignoring or naïvely replacing observations below the LOD, lead to negatively biased estimates of the AUC. Maximum likelihood methods have been developed for estimating the parameters of a bivariate normal distribution for two biomarkers with LODs. As well as assessing the negative effects of naïve replacement, we generalize the bivariate likelihood for left censored normal random variables to p biomarkers with LODs, develop point estimator and confidence interval for AUC and demonstrate asymptotic unbiasedness and nominal coverage probability, respectively.

Estimates of Parameters in the Incidence Model for the Two Hospital Capture-Recapture Problem

✱Lawrence Lessner, State University of New York at Albany, School of Public Health, Epi & biostat, One University Place, Rm 150, Rensselaer, NY 12144-3456, *LLessner@nycap.rr.com*

Key Words: capture-recapture, Poisson, marked, point process, Bayesian

This abstract is for a capture recapture model for the incidence of a given disease among patients in a given region: see session 34, July 29, 2007. The model uses marked, nonhomogeneous Poisson process to represent the incidence of a patient selecting hospital A, hospital B or neither A or B. The data maybe represented by a 2 by 2 table where one cell is not observed. Earlier work showed that the observations in each cell of the 2 by 2 table were independent Poisson random variables, and the assumption of independence in the "capture probabilities" was shown to be not true. This presentation will be devoted to an exploration of the maximum likelihood estimates obtained from the incidence model and the applications of these estimates. Also additional information is employed in a variety of Bayesian methods.

357 Data Mining and Machine Learning

Section on Statistical Computing

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Contiguous Clustering of Temporal Data

✱James A. Shine, U.S. Army Topographic Engineering Center, 4218 Alcott Street, Alexandria, VA 22309, *James.A.Shine@usace.army.mil*; James E. Gentle, George Mason University

Key Words: clustering, temporal, spatio-temporal, contiguous

There are many applications where dividing temporal data into two or more contiguous groups is valuable. For example, in stock data, it is useful to see how daily rates of return will cluster over a period of a year, to determine different types of markets. The authors have developed a basic algorithm to perform this clustering, and have several variations to allow different factors to vary in their influence. The algorithm has been tested and evaluated on several simulated and real-world data sets. We also have investigated similar contiguous clustering ideas for spatial and spatio-temporal data. Results and conclusions will be presented.

Improving Identification of the Minority Class on Imbalanced Data Sets

✱Chen-An Tsai, China Medical University, Biostatistics Center & Department of Public Health, Taichung, International 404 Taiwan, *catsai@mail.cmu.edu.tw*; James J. Chen, National Center for Toxicological Research

Key Words: class prediction, perturbation procedure, sensitivity, specificity, microarray

Class prediction is one of the most important problems in supervised learning. Many practical datasets often have a skewed class distribution in which at least one of classes has very small observations, usually the class of interest. As a result, a classifier induced by such an imbalanced data set appears to have typically high accuracy for majority classes and poor prediction for the minority class. In this article, we propose a perturbation procedure by adding observations with small disturbances to the minority class in order to squeeze out additional useful information. The Monte Carlo simulations show that the use of perturbation is useful to improve performance of either sensitivity or specificity of a classifier under specific conditions. In addition, the proposed approaches are applied to microarray data sets for illustrating the improvement of classifier performance.

The Ensemble Tree

✱J. Brian Gray, The University of Alabama, ISM Department, Box 870226, Tuscaloosa, AL 35487-0226, *bgray@cba.ua.edu*; Jie Xu, The University of Alabama

Key Words: bagging, boosting, decision tree, ensemble, predictive model, random forest

Decision trees are important tools in data mining with high interpretability and good predictive performance (Breiman 2001a). Ensemble techniques, including bagging (Breiman 1996), boosting (Freund and Schapire 1996, Friedman, Hastie, and Tibshirani 2000), and random forests (Breiman 2001b), have better predictive performance than decision trees, but this increased performance comes at the loss of interpretability (Breiman 2001a). In this paper, we propose the ensemble tree, a new predictive model that blends the accuracy of ensemble models and the interpretability of decision trees. The ensemble tree is illustrated on a variety of data examples and its performance is compared to that of decision trees and ensembles.

Generalized Reduced Error Logistic Regression Machine

✱Daniel M. Rice, Rice Analytics, 10805 Sunset Office Drive, Suite 300, St. Louis, MO 63127, *drice@riceanalytics.com*

Key Words: reduced error logistic regression, machine learning, SAS, penalized logistic regression, multicollinearity, overfitting

Reduced Error Logistic Regression (RELR) is a 100% automated method with no arbitrary or validation-sample parameters. RELR's error reduction results from symmetrical constraints consistent with Extreme Value properties of the Logit error. These constraints also lead to a prior ordering of the importance of variables, so the vast majority of variables can be excluded to avoid dimensional curse. RELR allows higher order polynomial terms and interactions to whatever order specified, but can give very parsimonious solutions with reasonable stability. Our M2007 paper showed that RELR can have better fit accuracy than SVM, DT, PLS, NN, and LR when multicollinearity is present. This paper compares RELR to Penalized Logistic Regression. We find that RELR can have better validation fit and solution stability without overfitting even though PLR uses the validation sample in its training.

Bonsai: Exploration and Cultivation of Machine Learning Models

✱David Purdy, University of California, Berkeley, 367 Evans Hall, #3860, Berkeley, CA 94720-3860, *dpurdy@stat.berkeley.edu*; Daisy (Zhe) Wang, University of California, Berkeley

Key Words: machine learning, visualization, data mining, decision trees, user interfaces, ensemble models

We introduce the Bonsai system for interactive visualization of high-dimensional machine learning models. We describe the system and how the design and usage is motivated to help users understand high dimensional machine learning models, the strengths and weaknesses of various models, and motivates designs of new algorithms tailored for the data set of interest. The system is configurable for most modeling methods, and we demonstrate applications involving CART, boosting, and SVMs, among others.

Borrowing Strength in Time Series Data Mining

✱Ganesh Subramaniam, AT&T Labs - Research, 180 Park Ave.,, Florham Park, NJ 07932, *mkg@research.att.com*; Ravi Varadhan, Johns Hopkins University

Key Words: Functional Data Analysis, Smoothing, Derivatives, Data Mining

In time series data mining the goal is look for interesting time series, (i.e., time series that deviate in interesting ways from typical response). There are several methods, one based on approach (a) using functional PCA due to Rice and Jones (1992) that help identify outlying time series. Another approach (b) Chaudhuri and Marron (1997) is to individually smooth curves and look at features of smoothed curves such as derivatives. Rice (2004) proposed the idea of borrowing strength commonly used in LDA (Diggle et al 1994) could be beneficially used in FDA settings. Here we explore whether borrowing strength gives us useful insights in time series data mining problems. We will explore this using a large telecom data and compare the results with approaches (a) and (b) in terms of its ability to provide additional insights and computational efficiency for data mining problems.

Data Complexity and Classifier Performance

✱Ana I. Landeros, The University of Alabama, 907 15th Ave Apt 2, Tuscaloosa, AL 35401, *alandero@cba.ua.edu*

Key Words: Data complexity, classifier performance

The performance of predictive models has been widely studied. Many evaluations have been conducted with considerable debate still existing as to which method provides the best approach. Different studies have been published reporting partial and/or contradictory outcomes. In general, these studies have not been conducted in a consistent manner so that the conclusions can be generalized beyond the specific research. To provide a better understanding of which predictive modeling technique works best and for what type of data we propose that is necessary to include measures to characterize the dataset complexity, since clearly, the performance of any classifier is data dependent. In this paper, we will introduce data complexity measures and their relation with model performance.

358 Financial and Economic Modeling

Section on Risk Analysis

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Percentile Estimators in Location-Scale Parameter Families Under Absolute Loss

✱Jerome Keating, University of Texas at San Antonio, Department of Demography, One UTSA Circle, San Antonio, TX 78249-0704, *jerome.keating@utsa.edu*; Robert L. Mason, Southwest Research Institute; Narayanaswamy Balakrishnan, McMaster University

Key Words: Equivariant Estimation, Median Unbiased, Absolute Risk, Percentiles

Equivariant estimators of percentiles of location-scale parameter families are optimized based on median unbiasedness and absolute risk. Median unbiased estimators and minimum absolute risk estimators are shown to exist within a class of equivariant estimators and depend upon medians of two completely specified distributions. These estimators are illustrated in the normal and exponential distributions.

Interval Estimation for the Largest Mean of a Multivariate Normal Population with Application to Mutual Fund Returns

✱Hubert J. Chen, National Cheng Kung University, 1 University Road, Tainan, Taiwan 701, Department of Statistics, Tainan, *hjpcheney@yahoo.com*

Key Words: Least Favorable, Optimal Interval, Correlated Populations, Mutual Funds, Inflation Rate

Assuming that there is a multivariate normal population with a common but unknown variance and common, nonegative correlation coefficients. The purpose is to find optimal confidence interval for the largest (and/or the smallest) mean of these correlated populations. The optimal confidence interval is obtained by maximizing the coverage probability at a least favorable configuration of means with the expected interval width being fixed. Some optimal properties are elaborated. Statistical tables of critical values are calculated to implement the proposed interval procedure. Finally, the interval is applied to the evaluation of the largest mean return (and the smallest one) of diversified mutual funds in the U.S. stock markets. It has shown, by our statistical inference, that long term investments in well-diversified mutual funds can almost always and largely beat the inflation rate.

Robust Neural Network with Application to Maturation Curve Estimation

✱Yijia Feng, The Pennsylvania State University, 326 Thomas Building, Department of Statistics, University Park, 16802, *yxf112@psu.edu*

Key Words: maturation curve, neural network, conditional quantile, nonparametric estimation

We study the robust neural network (RNN) for nonparametric conditional quantile estimation. We proposed an MM algorithm to realize the optimization of the general quantile loss. A simulation study is performed to compare the proposed RNN with some other nonparametric regression methods. An application of our method is presented to estimate the maturation curve in a credit card portforlio data set.

Risks Estimation in Economic Systems on the Basis of the Processing Approach

✱Dmitry G. Morgunov, Novosibirsk State Technical University, Cheluskintsev St. 2, flat 4, Novosibirsk-4, 630004 Russia, *dima_sw87@mail.ru*; Anatoly A. Naumov, Novosibirsk State Technical University

Key Words: Risk estimation, Processing approach, Managing economic systems, Conditional risks, Internal risks, External risks

The task of risk estimation and using it on practice in managing economic systems with using processing approach is investigated. Conditional-internal and conditional-external risks are considered and proved their properties. Metrics of risks estimation are given as highest comparative and absolute losses, average losses etc. Uneffecienty of using risk estimation procedures by dispersion and quantiles is shown. Approaches to the business-processes indicators optimization proposed and investigated in the area "risk-profit." This optimization is realized with using utility function as multicriterion problem and investigated in different posings. Functional, stochastic dependences between risks are shown, that do not allow to optimize risks of separate indicators. The task of risk managing in economic systems is work out in complex with the tasks of it's analysis, modelling and optmization.

359 Reliability and Censored Data Analysis

Section on Quality and Productivity, Section on Physical and Engineering Sciences

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Weekday Dependence in Reliability Analysis of Repairable Systems

✱Alexandre Zolotovitski, Sun Microsystems, Inc., San Diego, CA 92130, *alex.zolot@gmail.com*; David C. Trindade, Sun Microsystems, Inc.

Key Words: Reliability analysis, field data, mean cumulative function, repairable system

A useful way of representing the reliability of repairable systems is plotting the Mean Cumulative Function and the Recurrence Rate (RR) versus the age or calendar date in days of the observed systems. Usually this approach implicitly assumes that systems operate uniformly 24 hours per day and seven days per week, the so-called 24/7. We analyzed weekday dependence statistics of failures and found that frequency of failure in weekends were often significantly lower than in work days. More detailed analysis showed that all systems could be separated into two groups: "5d systems" that never had failure on Saturdays and Sundays and "7d systems" that had failures any day of week. Typically, the ratio of average weekly RR's for these two groups is close to 5:7. Neglecting this fact and aggregating statistics of systems of both types can lead to extra noise and bias in parameter estimation.

Estimation of Percentage of Defective Items in Quality Control

✱Shui-Ching Chang, Overseas Chinese Institute of Technology, Department of Information Management, 100 Chiao Kwang Rd., Taichung, 407 Taiwan, *monet@ocit.edu.tw*; Tze Fen Li, Ming Dao University

Key Words: generalized gamma distribution, concave function, life distribution, maximum likelihood estimation, mixture, prior probability

The generalized gamma distribution (g.g.d.) is often suggested as a life distribution because exponential, Weibull and gamma distributions are the special cases of the g.g.d.. It is difficult to assume a certain probability distribution for the life time. Hence the g.g.d. is widely used for life distribution and the random life time will find a suitable distribution within the g.g.d.. This paper is to find the defective percentage in an unclassified item for quality control. The results of a Monte Carlo simulation study with the g.g.d. (including Weibull, gamma and exponential distributions) are presented to show the accurate estimation of the defective percentages.

Mean Residual of Some Discrete Distributions

✱Syed A. Hossain, Rider University, Management Science Department, Lawrencevilleq, NJ 08648, *shossain@rider.edu*; Mohammad Ahsanullah, Rider University; Syed N. Kirmani, University of Northern Iowa

Key Words: Mean residual, Discrete distributions, Conditional expectation, Hazard rate, Characterization

We present some distributional properties of mean residual of discrete distributions. It has been found that for most of these distributions the mean residuals are related to their hazard rates. Using these relations several characterizations of the discrete distributions are presented. Results are useful in reliability and actuarial science.

Estimating Performance Degradation Using Intervals Between Upcrossings of a Threshold

✱Brock E. Osborn, GE Global Research, Applied Statistics Laboratory, Niskayuna, NY 12309, *osborn@research.ge.com*; Thomas R. Willemain, Rensselaer Polytechnic Institute; Hui Fan, Moody's Investors Services; Pasquale Sullo, Rensselaer Polytechnic Institute

Key Words: upcrossing, level crossing, censoring, reliability, degradation

We consider the problem of estimating the rate of degradation of a system with severely censored data. Specifically, we consider the problem of deriving a consistent estimator of degradation when we can only observe the times at with an unobserved performance measurement exceeds a given threshold. Such problems arise, for example, in the operation of complex systems where exceeding a predefined level multiple times drives repair actions, in security applications for which only entry into a secured facility is recorded, and in sensor networks where data acquisition and transmission is limited. We consider the relative efficiency of the resulting estimator compared to the MLE estimators derived from the full unobserved process and from unobserved binary high-low data.

A Method for Calculating Control Limits for Multiple Characteristics That Are Skewed and/or Detection Limit Censored

✱Robert Brill, ICL Performance Products, 532 Millstone Drive, Lawrence, KS 66049, *robert.brill@icl-pplp.com*; Thomas J. Bzik, Air Products and Chemicals, Inc.

Key Words: control limits, skewed data, detection limit censoring, family type I error, individual type I error, non-distributional

There is a trend in industry for producers to require their raw material suppliers to provide materials that not only meet nominal specifications but also that each characteristic be in statistical control. This has the affect of making control limits de facto specifications. In the semi conductor industry a single raw material typically has more than 20 specifications. The data required to calculate control limits are often highly skewed and censored to detection limits. The SEMI trade organization has recently issued a Guideline, SEMI C64-0308 - SEMI Statistical Guidelines for Ship To Control, for calculating control limits for such data that is non-distributional, does not require transformation of the data and approximately maintains a set type I error at the family level. This talk will explain the statistical background of the guideline.

Nonparametric Test Families for Highly Censored and/or Skewed Data

✱Thomas J. Bzik, Air Products and Chemicals, Inc., 5613 Tavern Circle, Macungie, PA 18062, *bziktj@airproducts.com*; Robert Brill, ICL Performance Products

Key Words: nonparametric, method detection limit, quantile, percentile, SEMI C64-0308, ship to control

A new industry standard, SEMI C64-0308 - SEMI Statistical Guidelines for Ship To Control, uses nonparametric methodologies to determine when underlying distributions have changed. These data distributions may be both highly skewed as well as highly censored to a method detection limit. A methodology for detecting distributional change by testing whether an upper or lower tail quantile has changed is provided. The 90th and 10th percentile case will be developed and the methodology can be extended to other percentiles of interest. An approximating equation that provides very reliable critical values for n1 and n2>=30 for the proposed test statistic is provided. A performance comparison is made to one and two-sided versions of Tukey's Quick Test. The new methodology provides a more powerful test.

Sequential Design for Analyses Involving Several Types of Data

✱Christine Anderson-Cook, Los Alamos National Laboratory, NM 87544, *c-and-cook@lanl.gov*; Todd Graves, Los Alamos National Laboratory; Michael Hamada, Los Alamos National Laboratory

Key Words: Resource Allocation, Reliability, Optimal Design

In analyzing the reliability of complex systems, several types of data from full-system tests to component level test are commonly used. After a preliminary analysis, additional resources may be available and we wish to identify the best new data to collect to maximally improve the prediction of system reliability. In this talk we present a simple example of a series system with multiple data sources, and show how to implement an optimization strategy to determine the best experimental design for subsequent data. Issues discussed including a metric for assessing improvement, discrepancies between data sources, incorporating cost into a comparison of different data types and implementation of an automated search for the optimal solution.

360 Estimation Methods in Selected Nonparametric and Semiparametric Settings

Section on Nonparametric Statistics, IMS
Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Estimating Monotone Convex Functions via Sequential Shape Modification

✱Sang Han Lee, Nathan Kline Institute, *shlee@nki.rfmh.org*; Johan Lim, Yonsei University; Seung-Jean Kim, Stanford University; Yongsung Joo, University of Florida

Key Words: Constrained least squares, convergence rate, monotone convexity, nonparametric estimation, uniform approximation

We propose a sequential method to estimate monotone convex function that consists of: (i) monotone regression via solving a constrained least square problem and (ii) convexication of the monotone regression estimate via solving an uniform approximation problem with associated constraints. We show that this method is faster than the constrained least squares (LS) method. The ratio of computation time increases as data size increases. Moreover, we show that, under an appropriate smoothness condition, the uniform convergence rate achieved by the proposed method is nearly comparable to the best achievable rate for a nonparametric estimate which ignores the shape constraint. Simulation studies show that our method is comparable to the constrained LS method in estimation error. We illustrate our method by analyzing ground water level data of wells in Korea.

Semiparametric Additive Isotonic Regression

✱Guang Cheng, Statistical and Applied Mathematical Sciences Institute, P.O. Box 14006, 19 T.W. Alexander Drive, Research Triangle Park, NC 27709-4006, *chengg@samsi.info*

Key Words: Isotonic Regression, Semiparametric Model, Additive Regression, Oracle Property

This paper is about the efficient estimation of the semiparametric additive isotonic regression model, i.e. $Y=X'\beta+\sum_{j=1}^{J}h_{j}(W_{j})+\epsilon$. Each additive component h_{j} is assumed to be a monotone function. It is shown that the least square estimator of the parametric component is asymptotically normal. Moreover, the isotonic estimator for each additive functional component is proved to have the oracle property, which means it can be estimated with the highest asymptotic accuracy, equivalently, as if the other components were known.

ANOVA for Semiparametric Models with Missing and Unbalanced Longitudinal Data

✱Pingshou Zhong, Iowa State University, 210 Snedecor Hall, Iowa State University, Ames, IA 50010, *pszhong@iastate.edu*; Song Xi Chen, Iowa State University

Key Words: ANOVA, Longitudinal data, Empirical likelihood

We consider ANOVA analyses for a semi-parametric partially linear regression model with missing and unbalanced longitudinal data from different treatment groups. We assume the missing is monotone and the propensity function is known up to a parameter. Empirical likelihood ratio statistics are formulated for both the equality of the parametric regression parameters and the nonparametric regression functions. It is found that the asymptotic distributions of these statistics, respectively, are chi-square and normal distribution, which lead to readily implementable nonparametric likelihood ratio ANOVA tests for both the parametric and nonparametric parts of the regression models. The performance of the ANOVA tests are evaluated empirically by simulation and a case study.

On the Asymptotics of Additive Penalized Splines

✱Yingxing Li, Cornell University, Ithaca, NY , *yl377@cornell.edu*; David Ruppert, Cornell University

Key Words: Asymptotics, P-spline, backfitting, penalty

In this talk, we study the asymptotic property when the p-dimensional additive model is fitted by penalized splines. The penalty is placed on the differences of the coefficients and B-splines are used. The penalized spline estimators are shown to have the same asymptotic distribution as that of backfitting using Nadaraya-Watson kernel estimators. The optimal convergence rate depends mainly on the order of the penalty given the number of knots exceeds a minimum bound. The asymptotic bias and variance are computed. Penalized spline estimators are not design-adaptive.

Parameter Estimation Under a Two-Sample Semiparametric Model

✱Jingjing Wu, University of Calgary, Department of Mathematics and Statistics, 2500 University Drive N.W., Calgary, AB T2N 1N4 Canada, *jinwu@math.ucalgary.ca*

Key Words: semiparametric model, Hellinger distance, kernel estimator, asymptotic normality

We investigate estimation problem of parameters in a two-sample semiparametric model, where the log ratio of the two underlying density functions is of a regression form. This model has wide applications in the logistic discriminant analysis, case-control studies, and receiver operating characteristic curves analysis. Furthermore, it can be considered as a biased sampling model with weight function depending on unknown parameters. In this paper, we construct minimum Hellinger distance estimators of the regression parameters. The proposed estimators are chosen to minimize the Hellinger distance between a semiparametric model and a nonparametric density estimator. Theoretical properties such as the existence, strong consistency and asymptotic normality are investigated. Robustness of proposed estimators is also examined using a Monte Carlo study.

Revisiting Qualms About Bootstrap Confidence Intervals

✱Michael R. Chernick, Smith Hanley Consulting Group, 15 Quail Drive, Holland, PA 18966, *mchernic@yahoo.com*

Key Words: bootstrap, confidence intervals, nonparametrics

In "Qualms about bootstrap confidence intervals," Schenker (1985) showed by example that the BC bootstrap breaks down for practical sample sizes. Efron devised the BCa method in "Better bootstrap confidence intervals," Efron (1987) to overcome the problem. We revisit the issues by estimating variances in a nonparametric setting. We show through simulation that even for a sample size n>=200 for certain heavy-tailed and skewed distributions, the convergence of the second order accurate bootstrap methods BCa, ABC, and tilted bootstrap is slow and all of the methods provide true confidence levels that are at least 5% below their advertised confidence levels. We illustrate this at confidence levels of 50%, 75% and 90%. For sample sizes n>=1000 to adequately show the proper pattern of convergence, we need the standard deviation of the Monte Carlo approximation of the proportion to be 0.005.

361 Developments in Nonparametric and Robust Time Series Analysis ▲

Section on Nonparametric Statistics, IMS

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

The Kernel Self-Normalized, Tail-Trimmed Sum for Dependent, Heterogeneous Data, with an Application to Robust Least Squares

✱Jonathan B. Hill, The University of North Carolina at Chapel Hill, Chapel Hill, NC , *jbhill@email.unc.edu*

Key Words: tail trimmed sum, Near Epoch Dependence, mixing, heavy tails, robust estimation, kernel variance

Altough robust estimation methods were formalized by the late 19th century, data trimming and truncation for non-iid data has received little attention. We establish sufficient conditions for asymptotic normality of a tail-trimmed sum of dependent, heterogeneous data. The sum is self-standardized with a kernel variance estimator so the rate of convergence, tail thickness and memory persistence do not need to be specified. The resulting central limit theory applies to mixing, geometrically ergodic, and Near-Epoch-Dependent processes in general, including linear and nonlinear distributed lags, FIGARCH and stochastic volatility processes, all with short or long memory and thin or thick-tailed shocks. The theory is applied to asymptotically Gaussian least squares estimation of an infinite variance first-order difference equation.

Analyzing Bivariate Time Series via Nonparametric Likelihood

✱Suddhasatta Acharyya, Brown University, Center for Statistical Sciences, 121 South Main Street, 7th Floor, Providence, RI 02912, *acharyya@stat.brown.edu*; Hernando Ombao, Brown University

Key Words: nonparametric-likelihood, cross-spectrum, coherence, time-series, brain-signal

Cross-spectrum and coherence are widely used measures for characterizing linear dependence between two time series. Inference on these quantities relies on the assumption of asymptotic Gaussian distributions. Such an assumption could be problematic when the length of the time series is not sufficiently large and when the noise innovations come from distributions with heavy tails. Nonparametric likelihood-based methods have provided competitive alternatives to standard parametric approaches for many problems. The application of such methods to the spectral domain has not been well-studied. In this paper, we shall assess the performance of nonparametric likelihood based estimation for spectral time series through simulation. We shall also apply the method to multi-channel brain signal data, and discuss possible implications.

Consistent Nonparametric Tests for Granger Causality

Yoshihiko Nishiyama, Kyoto University, Institute of Economic Research; Kohtaro Hitomi, Kyoto Institute of Technology; ✱Yoshinori Kawasaki, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo, International 106-8569 Japan, *kawasaki@ism.ac.jp*; Kiho Jeong, Kyungpook National University

Key Words: time series, nonlinear causality, nonparametric test, local alternatives

Most of the existing tests of Granger-type causality are concerned only with "linear causality in mean," namely when a series linearly affects the (conditional) mean of the other series. However, dependence between series may be nonlinear, and/or not only through the conditional mean. The purpose of this paper is to propose nonparametric tests for possible nonlinear causality up to K-th conditional moment. A desirable property of the tests is that they have nontrivial power against square-root-T-local alternatives, where T is sample size. Their null asymptotic distributions are not normal, but we can easily calculate the critical regions by simulation. Monte Carlo experiments show that the proposed tests have good power properties, and much better than the competitors. An application is given in the context of causality in conditional heteroscedasticity.

Spline-Backfitted Additive Nonparametric Transfer Function Models

✱Jun Liu, Georgia Southern University, PO Box 8151, Statesboro, GA 30460-8151, *jliu@georgiasouthern.edu*

Key Words: nonparametric, transfer function, backfitting, additive model

In this paper we consider additive modeling of multi-dimensional nonlinear transfer functions with ARMA type of noise. Polynomial splines are used to obtain preliminary estimates of the additive transfer function components jointly with the ARMA parameters. These preliminary estimates are then used in backfitting to obtain final estimates of the transfer function components, using local polynomial regression. By showing that the errors caused by spline approximation in the preliminary estimation is negligible, it can be shown that an additive component can be estimated by local polynomial asymptotically as if the other components and the ARMA parameters are known. This spline-backfitted estimator is less computationally intensive than a full-blown local polynomial additive model, and its asymptotic properties are easier to derive than a full-blown polynomial spline additive model.

Kernel Reweighting for Inference on Time Series

✱Kristofer Jennings, Purdue University, Department of Statistics, 250 North University Street, West Lafayette, IN 47907-2066, *jennings@stat.purdue.edu*

Key Words: block bootstrap, tapering, random weighting bootstrap

Many bootstrap-based methods for time series inference involve the construction of a “pseudo series” of blocks of the original observations. The kernel reweighting procedure uses a correlated reweighting of the original statistic that does not rely on having a reconstructed series. As such, it resembles the tapered block bootstrap of Paparoditis and Politis (2001) without restrictions on the size of tapering window. The asymptotic properties are similar to those of the tapered block bootstrap for stationary series. However, the flexibility of reweighting allows for asymptotic improvement via iteration. A criterion for choosing block size is given. An empirical likelihood formulation is also discussed.

Testing for the Equality of Two Autoregressive Functions Using Quasi-Residuals

✱Fang Li, Indiana University Purdue University, Indianapolis, Department of Mathematical Sciences, 402 N. Blackford Street, LD270, Indianapolis, IN 46202, *fli@math.iupui.edu*

Key Words: autoregressive, nonparametric, quasi-residual, kernel estimator

This paper discusses the problem of testing the equality of two nonparametric autoregressive functions against one-sided alternatives. The heteroscedastic errors and stationary densities of the two independent strong mixing strictly stationary time series can be possibly different. The paper adapts the idea of using sum of quasi-residuals to construct the test and derives its asymptotic null distributions. The paper also shows that the test is consistent for general alternatives and obtains its limiting distributions under a sequence of local alternatives. Then a Monte Carlo simulation is conducted to study the finite sample level and power behavior of these tests at some alternatives. We also compare the test to an existing lag matched test theoretically and by Monte Carlo experiments.

Local Linear Quantile Estimation for Nonstationary Time Series

✱Zhou Zhou, The University of Chicago, Department of Statistics, 5734 S. University Ave., Chicago, IL 60637, *zhouzhou@uchicago.edu*

Key Words: quantile estimation, local stationarity, non-stationary nonlinear time series, climate change

We consider estimation of quantile curves for a general class of non-stationary processes. Consistency and central limit results are obtained for local linear quantile estimates under a mild short-range dependence condition. Our results are applied to environmental data sets. In particular, our results can be used to address the problem of whether climate variability has changed, an important problem raised by IPCC (Intergovernmental Panel on Climate Change) in 2001.

362 Bayesian Design and Bayesian Methods for Poisson Data ●▲

Section on Bayesian Statistical Science

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

The Impacts of Misclassification on Bayesian Adaptive Designs

✱Jo A. Edmonds, The University of Kansas Medical Center, 4816 Horton St., Mission, KS 66202, *jedmonds2@kumc.edu*; John W. Seaman, Jr., Baylor University; James Stamey, Baylor University

Key Words: Bayesian Statistics, Adaptive Design, Misclassification, Clinical Trials

Among the methods that have become increasingly important in drug development are adaptive experimental designs. A primary argument for the use of adaptive designs is the efficiency one gains over implementing a traditional fixed design. We consider a simple two-arm Bayesian adaptive design utilizing adaptive allocation in which the binary outcome is subject to misclassification. A model is developed to incorporate the misclassification in the response, prior specification and issues with convergence are discussed, and we conclude with the results of a simulation study performed to assess the impact of misclassification on the efficiency and performance of the design.

Exact Bayesian Inference in Matched Pairs Designs

✱Yong Chen, Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, JHSPH, 615 North Wolfe Street, Baltimore, MD 21205, *yonchen@jhsph.edu*; Sining Chen, Johns Hopkins University; Haitao Chu, The University of North Carolina at Chapel Hill

Key Words: Matched Pairs Designs, Exact Bayesian Inference

Matched binary pairs arise frequently in biomedical research including clinical trials involving pre-/post comparison, and observational studies such as matched case-control and matched cohort studies. The relative risk, odds ratio and risk difference are commonly used as the measurements of the association between exposure and outcome in matched binary pairs. However, to our knowledge, the exact probability distributions of those measures have not been given in the literature. In this paper, we derive the exact posterior distributions of these measures for paired binary endpoints, which could be widely used in biomedical research. Furthermore, we discussed Bayesian inference based on these exact distributions, particularly when the sample size is small. The proposed method is applied to several real biomedical studies.

Bayesian Optimal Single Arrays for Robust Parameter Design

✱Lulu Kang, Georgia Institute of Technology, School of Industrial and Systems Engineering, 765 Ferst Drive, NW, Atlanta, GA 30332, *lkang@isye.gatech.edu*; Roshan J. Vengazhiyil, Georgia Institute of Technology

Key Words: Design of experiment, Exchange algorithms, Quality improvement, Variation reduction

We suggest a new optimal design criterion for robust parameter design experiment. This criterion is built on a Bayesian framework which has incorporated the hierarchical ordering principle. Compared with some existing design criteria, this criterion still follows the hierarchical ordering principle and fully focuses on variation reduction. We also develop a greedy-exchange algorithm to search for the optimal design. Our proposed method is very general and not restricted to a specific kind of designs such as orthogonal arrays. It can be extended to more complicated situations when there are mixed-level qualitative and quantitative factors, or even internal noise factors.

Comparison of the Bayesian Prediction Limits for the Poisson Process

✱Valbona Bejleri, University of the District of Columbia, 4200 Connecticut Avenue, N.W., Washington, DC 20008, *vbejleri@udc.edu*; Alexander White, Texas State University, San Marcos

Key Words: Prediction, Poisson, Bayesian, Non informative, Frequentist

Bayesian limits for the Poisson distribution with informative and noninformative priors are compared based on their frequentist properties. For any prior other than the modified Jeffreys' prior, it is shown the Bayesian lower prediction limit derived from it is either smaller or the frequency of making a wrong prediction exceeds the probability of wrong coverage. Hence, the lower prediction limit derived from Jeffreys' prior is the only Bayesian lower limit that coincides with the frequentist lower limit. It is shown that this is not the case for the upper prediction limit derived from Jeffreys prior; the upper prediction limit does not always coincide with the frequentist upper prediction limit. Analog results are obtained in the case of uniform prior. It is shown there is no prior distribution such that the upper and lower Bayesian limits always coincide with the frequentist limits.

Characterizing the Performance of a Bayesian Conway-Maxwell GLM

✱Seth D. Guikema, Johns Hopkins University, Department of Geography and Environmental Engineer, Baltimore, MD 21218, *sguikema@jhu.edu*; Srinivas Geedipally, Texas A&M University; Dominique Lord, Texas A&M University; Soma Dhavala, Texas A&M University

Key Words: generalized linear model, Conway-Maxwell Poisson, Count data regresssion, MCMC

The Conway-Maxwell Poisson (COM) distribution is an extension of the Poisson distribution that can model both over and under dispersion. Previous work by the authors developed a dual-link generalized linear model based on the COM distribution together with a MCMC implementation of the model and applied this model to analyzing power system reliability and traffic accidents. In this paper we characterize the performance of this COM GLM and its MCMC implementation in terms of both parameter estimation accuracy and computational burden. Using simulated data, we show that the COM GLM can handle under, equi, and overdispersed data sets with low, moderate, and high means. We show that the MCMC implementation of the COM GLM yields accurate parameter estimates and that the computational burden of the COM GLM is not prohibitive. The COM GLM offers a flexible alternative for count data regression.

Bayesian Detection of Changes of a Poisson Process Monitored at Discrete Time Points Where the Arrival Rates Are Unknown

✱Marlo Brown, Niagara University, Dunleavy Hall, Niagara University, NY 14109, *mbrown@niagara.edu*

Key Words: Bayesian stopping rules, Change-point detection, Dynamic Programming, Poisson processes, Risk, Unknown arrival rates

We look at a Poisson process where the arrival rate changes at some unknown integer. At each integer, we count the number of arrivals that happened in that time interval. We assume that the arrival rates before and after the change are unknown. For a loss function consisting of the cost of late detection and a penalty for early stopping, we develop, using dynamic programming, the one and two steps look ahead Bayesian stopping rules. We provide some numerical results to illustrate the effectiveness of the detection procedures.

Estimation of Web Page Change Rates

✱Carrie Grimes, Google, Inc., 1600 Amphitheatre Pkwy, Mountain View, CA 94043, *cgrimes@google.com*; Daniel Ford, Google, Inc.

Key Words: changerate estimation, bayesian estimation, poisson process models

Search engines strive to maintain a "current" repository of all pages on the web to index for user queries. However, crawling all pages all the time is costly and inefficient: many small websites don't support that much load and while some pages change very rapidly others don't change at all. Therefore, estimated frequency of change is often used to decide how often to crawl a page. Here we consider the effectiveness of a Poisson process model for the updates of a page, and the associated Maximum Likelihood Estimator, in a practical setting where new pages are continuously added to the set of rates to be estimated. We demonstrate that applying a prior to pages can significantly improve estimator performance for newly acquired pages.

363 Spatial Modeling and Related Computational Issues

Biometrics Section, Section on Statistical Computing

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Joint Spatial Modeling of Recurrent Tree Infection and Growth with Processes Under Intermittent Observation

✱Farouk S. Nathoo, University of Victoria, PO Box 3045 STN CSC, Mathematics and Statistics, Victoria, BC V8W3P4 Canada, *nathoo@math.uvic.ca*

Key Words: recurrent events, panel data, growth curves, multivariate spatial process, kernel convolution, joint modelling

In this article we present new methods for longitudinal studies in forestry where trees are subject to recurrent infection and the hazard of infection depends on tree growth over time. Here, tree height is an intermittently observed, time-dependent and spatially-varying covariate that is subject to measurement error. We develop a joint model linking a counting process for recurrent infections with a nonlinear spatial growth curve representing the underlying height trajectories. To accommodate spatial variability in growth parameters corresponding to trees at different locations, a flexible and robust non-Gaussian multivariate spatial process model is developed. We adopt an approach based on kernel convolution with the kernel modeled in a semi-parametric fashion. In particular, a penalized smoothing approach based on a piecewise constant kernel and Markov random field priors is employed.

An Approach to the Analysis of Spatially Correlated Multilevel Functional Data

✱Ana-Maria Staicu, University of Bristol, Department of Statistics, University Walk, Clifton, Bristol, International BS8 1TW United Kingdom, *a.staicu@bristol.ac.uk*; Ciprian M. Crainiceanu, Johns Hopkins University; Raymond Carroll, Texas A&M University

Key Words: multilevel functional data, spatial correlation, principal compoment analysis, kernel estimation

We describe the framework and inferential tools for hierarchical functional data where the functions at the lowest hierarchy level are spatially correlated. We present a new approach in which the hierarchical functions are modeled nonparametrically using multilevel eigenfunction bases, which appear in a multilevel functional principal component scenario, plus a weakly stationary process to account for the spatial correlation. For each level, the eigenfunction basis is estimated from the data and the functional principal component score estimates are obtained by a conditional step; a method which is conceptually simple and straightforward to implement. A second novelty of our methodology is in using kernel smoothing estimation to estimate the spatial covariance function. The proposed procedure is illustrated with a simulation study and p27 measurements data in a carcinogenesis study.

Bayesian Spatial Analysis of Sexual Behaviors Among HIV High-Risk Individuals in Sexual Networks

✱Robin Jeffries, University of California, Los Angeles, 3220 Midvale Ave., School of Public Health, Department of Biostatistics, Los Angeles, CA 90034, *rjeffries@ucla.edu*; Robert E. Weiss, University of California, Los Angeles; Pamina Gorbach, BERG; Steve Shoptaw, University of California, Los Angeles

Key Words: autoregressive spatial prior, hierarchical model, random effects, repeated measures

We develop Bayesian statistical methodology to analyze multivariate self-reported sexual behaviors in a sample of injecting drug users and men who have sex with men. Subjects were sampled by respondent driven sampling: current subjects recruit the next wave of subjects from among their sexual partners and people they use drugs with. This induces correlation between subjects and their recruits but provides information on the sexual and drug-using networks of people at risk for HIV. Risk behavior information is collected on up to three sex and three drug partners for each individual surveyed, giving correlated repeated measures data on subjects in a network that we model with a multivariate spatial random effects model. By jointly analyzing unprotected and protected anal intercourse, we can determine how drugs affect risky behaviors and how safe people are with different partner types.

Fitting Regression Models of Propagation Type to the Boolean Random Set

✱Khalil Shafie, University of Northen Colorado, Applied Statistics and Research Methods, 532 McKee Hall Campus Box 124, Greeley, CO 80639, *khalil.shafie@unco.edu*; Mojtaba Khazaee, Shahid Beheshti University; Mojtaba Ganjali, Shahid Beheshti University

Key Words: Random set, Hitting functional, Generalized linear model, Boolean model, Propagation model

In this work the regression problem for random sets of the Boolean model type is developed, where the corresponding Poisson process of the model is related to some explanatory variables, assuming the random grains are not affected by these variables. The propagation model is introduced and some methods for fitting this model are presented. Then the model is applied in a simulation study.

Why Do Ticks Like My Dog?

✱Hyun-Joo Kim, Truman State University, 100 normal, kirksville, MO 63501, *hjkim@truman.edu*; Stephanie A. Fore, Truman State University; Laura J. Fielden-Rechav, Truman State University; Georgia Muller, Truman State University; Tad Dallas, Truman State University; Will Petry, Truman State University

Key Words: Ticks, Parasite, Home range, Hostseeking behavior, Web and Grid comparison

A parasite is defined as "an organism that lives in or on a host from which it derives its food and other biological supplies" (Kim 1985). Host-parasite associations often reflect a long history of coevolution between the two and may range from generalist to highly host specific. Rodents and its parasite, particularly ticks, are collected in northern Missouri. In this paper we will discuss about the home range of the rodents and its relationship with the parasite loads. Tick's host-seeking behavior will also be discussed.

Bayesian Mixture Labeling by Posterior Modes

✱Weixin Yao, Kansas State University, 101 Dickens Hall, Kansas State Univeristy, Manhattan, KS 66506, *wxyao@ksu.edu*; Bruce G. Lindsay, The Pennsylvania State University

Key Words: Bayesian mixtures, Label switching, Mixture models

One of the most fundamental problems for Bayesian mixture model analysis is label switching, which occurs due to the non-identifiability of the mixture components under symmetric priors. The widely used and easiest way to solve the label switching is to use an explicit parameter constraint so that only one permutation can satisfy it. Here, we propose a new labeling method we call PM(ECM), which assigns the labels based on the modes one converges to using the ECM algorithm. Due to the ascending property of the ECM algorithm, the PM(ECM) labeling method automatically matches the "ideal" labels in the highest posterior density (HPD) credible region. Hence the PM(ECM) creates a natural and intuitive partition of the parameter space into labeled regions. Using a real data set and a Monte Carlo simulation study, we demonstrate how effectively our new labeling method works.

Numerical Identifiability and Convergence for Gibbs Sampling of Nonlinear Models

✱Seongho Kim, Indiana University School of Medicine, Division of Biostatistics, Department of Medicine, 410 West 10th Street, Suite 3000, Indianapolis, IN 46202, *kim203@iupui.edu*; Lang Li, Indiana University

Key Words: Convergence, Gibbs Sampler, Nonidentifiability

Although formal identifiablility is not an issue in a Bayesian viewpoint, a poorly behaved Gibbs sampler frequently arises. The relationship between the prior selection and the convergence rate of a Gibbs sampler has been mostly investigated on nonidentifiability. The objective of this article, however, is to investigate the relationship between two Gibbs sampling schemes and their convergence rates on the numerical identifiability of parameters of a nonlinear model, given a prior information. We clarify that separate components in a Gibbs sampler result in a more efficient sampling scheme on nonidentifiability, while grouping random components are more efficient on identifiability. We also suggest a scheme for checking numerical identifiability of the parameters at the initial estimates for given experiments on the model.

364 Econometric Time Series

Business and Economics Statistics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Test of Cointegration Using Long-Run Canonical Correlations: Part II

✱Kalidas Jana, University of Texas at Brownsville, Department of Business Administration, 80 For Brown, Brownsville, TX 78520, *kalidas.jana@utb.edu*

Key Words: Long Run Canonical Correlations, Cointegration, Squared Simple Coherence at Frequency Zero, Squared Multiple Coherence at Frequency Zero, Squared Canonical Coherences at Frequency Zero

Jana (2007) in Part I investigated implications of cointegration for long run canonical correlations and equivalently, for canonical coherences at frequency zero. In addition, employing existing tests of zero squared simple, multiple, and canonical coherences at frequency zero, this part proposed new tests of null hypothesis of no cointegration. In Part II we conduct simulation studies to judge the performance of these proposed tests.

A Decomposition Approach to Revisions in Aggregated Time Series Estimates

✱Duncan J. Elliott, Office for National Statistics, Cardiff Road, Newport, International NP10 8XG United Kingdom, *Duncan.Elliott@ons.gov.uk*; Craig H. McLaren, Office for National Statistics; Julian Chow, Office for National Statistics

Key Words: revisions, seasonal adjustment, aggregation, time series, decomposition

Users of official statistics are often concerned with the source and impact of revisions in time series estimates. In official statistics it can be difficult to assess the impact and source of revisions where seasonally adjusted estimates are derived indirectly through the aggregation of seasonally adjusted component series. We investigate how higher level aggregates are impacted by revisions from directly seasonally adjusted time series by using a decomposition framework. The implications of this approach for interpretation of time series estimates and the production process in official statistics are also discussed.

Forecasting Methods for GDP with Measurable Risks

✱Les Yen, University of Phoenix/NVA, 11730 Plaza America Drive, Suite 200, Reston, VA 20190, *statexp@yahoo.com*; Eric Cramer, University of Phoenix

Key Words: GDP, Forecasting methods, Risk analysis, Nonstationary processes, Monte-Carlo simulations, ARIMA

Typically, the Gross Domestic Product (GDP) follows a long-run trend, where temporary contractions affect its long term level. Research economists have developed highly technical forecasting methods to assess the GDP secular trends using multiple regression models. These regression-type models were based on data observed sequentially over a stationary time series and were almost all unsuccessful in predicting, for example, the recession of the 1970s and 1990s. This paper examines and evaluates various forecasting methods for GDP based on nonstationary data and probabilistic risk analysis. We will consider a class of nonstationary series represented by ARIMA processes that can be reduced to ARMA processes. We will rely on Monte-Carlo methods to generate long-run simulated events and determine the likelihood of the risk occurring.

From California to Connecticut: Examining House Price Models for the USA

✱Chaitra Nagaraja, The Wharton School, University of Pennsylvania, *chaitra@wharton.upenn.edu*; Lawrence Brown, University of Pennsylvania

Key Words: housing, autoregressive, repeat sales

Models to predict house prices come in all flavors. In the repeat sales setup, homes that have been sold only once are omitted from the analysis. However, these single sales often constitute a substantial proportion of the data. The proposed repeat sales model eliminates this weakness. The model is composed of a fixed time effect combined with an autoregressive component. The latter piece is applied only to homes sold repeatedly while the former component utilizes all of the data. To evaluate the proposed model, single-family home sales for twenty U.S. metropolitan areas regions are analyzed. The proposed model is shown to have better predictive abilities than the established S&P/Case-Shiller model.

Estimation and Inference in Unstable Models Estimated via Nonlinear Least Squares

✱Otilia Boldea, Tilburg University, Warandelaan 2, Tilburg, International 5038VW The Netherlands, *otilia.boldea@uvt.nl*; Alastair Hall, University of Manchester and NCSU

Key Words: Structural Change, Multiple Break Points, Nonlinear Least Squares, Nonlinear Asymptotics, Empirical Processes

In this paper, we extend Bai and Perron's (1998, Econometrica, pp. 47–78) method for detecting multiple breaks to nonlinear models. To that end, we consider an unstable univariate nonlinear least squares (NLS) model with a limited number of parameter shifts occurring at unknown dates. In our framework, the break-dates are simultaneously estimated with the parameters via minimization of the residual sum of squares. Using nonlinear asymptotic theory, we derive the asymptotic distributions of both break-point and parameter estimates and propose several instability tests. We also present simulation results that validate our procedure. Our method is useful for estimating and testing nonlinear macroeconomic models with multiple unknown breaks. By construction, it also removes the usual confusion between nonlinearities and breaks.

Nonlinear Dynamics and Persistence in PPP Relation: Does Controlling for Nonlinearity Solve the PPP Puzzle?

✱Rehim Kilic, Georgia Institute of Technology, *rehim.kilic@econ.gatech.edu*

Key Words: PPP, nonlinearity, persistence, regime-dependent half-life

We use half-life measures conditional on various PPP regimes, to examine dynamic evolution of persistence in deviations from PPP within the context of ESTAR models. We present both asymptotic and simulated confidence intervals. Small sample performance and properties of regime-dependent half lives as the nonlinear dynamics and the persistence evolves in an ESTAR model are examined by Monte Carlo experiments. Results show that the precision of estimated persistence increases with the nonlinearity. Reported empirical results suggest both fast and slow reversion in PPP as well as noticeably different persistence dynamics across Euro and non-Euro zone currencies. Confidence bounds reveal notable uncertainty and persistence in several quarterly U.S. Dollar PPP deviations over the floating period, hence controlling for nonlinearity in PPP may not necessarily resolve the PPP puzzle.

365 Proteomics and Metabolomics

Biometrics Section, Section on Nonparametric Statistics, WNAR

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

A Systematic Low-Level Analysis for Both SELDI- and MALDI-TF-MS Data with Improved Peak Alignment Algorithm

✱Xuelian Wei, University of California, Los Angeles, 234 Escuela Ave, Apt 32, Mountain View, CA 94040, *wxl@stat.ucla.edu*; Kerchau Li, University of California, Los Angeles; James LeBlanc, University of California, Los Angeles; Thomas Drake, University of California, Los Angeles

Key Words: SELDI-TF-MS, MALDI-TF-MS, peak alignment, low-level analysis, mass spectrometry

In this study, we first summarized the pre-process steps for mass spectrometry analysis proposed recently and provided a systematical pre-process procedure suitable for both SELDI- and MALDI-TF-MS data. And then we focus on the peak alignment step and proposed so-called correlation-based hierarchical clustering (CBHC). This new method is motivated form the complete linkage hierarchical cluster (CLHC), but we add many important modifications to best ultimate all the information from both the locations and the shapes of detected peaks to form a better peak alignment. We validate our method on two real mass spectrometry datasets. We show that CBHC produces more reasonable peak alignment than CLHC and the protein profiles detected by CBHC have lower variations among replicated spectra and help to increase the sensitivity of follow-up biomarker detection comparing with CLHC and Ciphergen.

Use of Left-Truncated Regression Models for Normalization of Global Mass Spectrometry Data

✱Douglas W. Mahoney, Mayo Clinic College of Medicine, 201 1st St. SW, Rochester, MN 55905, *mahoney@mayo.edu*; Christopher J. Malone, Winona State University; Ann L. Oberg, Mayo Clinic College of Medicine; Jeanette E. Eckel-Passow, Mayo Clinic College of Medicine; Terry M. Therneau, Mayo Clinic College of Medicine; H. Robert Bergen, III, Mayo Clinic College of Medicine

Key Words: Mass Spectrometry, Normalization, Censoring, iTRAQ

Shotgun proteomics via mass spectrometry is a powerful technology for study of the proteome; it has the potential to lead to a noninvasive screening mechanism of proteins in easily accessible body fluids. With the iTRAQ labeling platform, four samples can be analyzed simultaneously with mass spectrometry. Although this increases through put and minimizes experimental error, there remains a need for normalization to validly compare two or more iTRAQ experiments. Due to the data-dependent acquisition process, there is a severe imbalance in the final data matrix making simple extensions of standard microarray normalization methods non-trivial. Through simulations, left truncated censoring models have been shown to adequately capture the normalization effects. We will compare censored and noncensored models as applied to iTRAQ data for normalization and differential expression analyses.

A New Algorithm To Solve Protein Identification Problems in MS/MS Analysis

✱Gabrila Cohen Freue, University of British Columbia, 333-6356 Agricultural Road, Vancouver, BC V6T 1Z4 Canada, *gcohen@mrl.ubc.ca*; Axel Bergman, Jack Bell Research Centre; Zsuzsanna Hollander, iCAPTURE; Robert Balshaw, iCAPTURE; Janet Wilson-McManus, iCAPTURE; Paul Keown, iCAPTURE; Bruce McManus, iCAPTURE; Robert McMaster, Jack Bell Research Centre; Raymond Ng, University of British Columbia

Key Words: protein identification, peptide list, protein groups, Mascot

Recently, iTRAQ in combination with Mascot gained in popularity among other technologies for protein identification and quantitation. Each protein is identified using a list of identified peptides. Usually, more than one protein is associated with each peptide list creating a protein group. Mascot selects one protein in each group for the protein summary table which contains the protein relative expressions. Thus, some proteins may appear to be missing for certain processed samples when they were present and identified under a different name. To maximize the number of proteins analyzed in a biomarker study, we proposed an algorithm that assigns an identification code to all the proteins in the same group. This code is then used to match proteins across different processed samples. The performance of this algorithm is illustrated in the analysis of protein data from transplanted patients.

Multidimensional Biomolecular NMR Studies: Noise Reduction and Component Identification

✱Nicoleta Serban, Georgia Institute of Technology, 765 Ferst Drive, NW, 438, Atlanta, GA 30318, *nserban@isye.gatech.edu*

Key Words: multi-dimensional NMR biomolecular studies, multi-dimensional wavelet-based noise reduction, multi-components identification

The overarching objective of the research presented at JSM is to shed light into the potential of wavelet-based methodology applied to data interpretation in NMR biomolecular studies. Specifically, the emphasis is on enhanced identification of spectral components in high-dimensional NMR frequency data using wavelet-based noise reduction. Our empirical investigation supports the benefits of initially reducing the noise for more detectable components and fewer false positives. The noise reduction method addresses a series of difficulties in NMR frequency data: local correlated noise, spatial heterogeneity and signal sparsity. The methodology is general, and therefore, it applies to other studies involving noise reduction of multidimensional data.

Empirical Bayes Methods for Biomarker Identification in Metabolomics

✱Cheng Zheng, Purdue University, Department of Statistics, IN 47906, *zhengc@purdue.edu*; Olga Vitek, Purdue University; Haiwei Gu, Purdue University; Zhengzheng Pan, Roche Diagnostics Co.; Daniel Raftery, Purdue University

Key Words: Mass Spectrometry, Direct Analysis in Real Time (DART), Desorption Electrospray Ionization (DESI), empirical Bayes, Linear Model for Microarray Analysis (LIMMA), metabolomics

Mass spectrometry-based metabolomics studies measure the abundance of metabolites in a sample at hundreds of mass to charge ratios (m/z). Determination of features that are differentially abundant between cases and controls is the primary goal. Empirical Bayes approaches (e.g., LIMMA) efficiently determine such features by assuming a common prior distribution of feature variances. However, the assumption is inappropriate for ionization techniques such as DART and DESI which produce a relationship between the features m/z and its variance. We extend the model in LIMMA by relating feature variances to m/z by a spline, and integrate the estimation into the empirical Bayes framework. We show by simulation that the proposed approach outperforms t-test, LIMMA, and EBarrays in terms of FDR and power. We illustrate the method on a metabolomic profiling experiment of breast cancer.

Statistical Analysis of Censored Proteomic Expression Data

✱Shiying Wu, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709, *swu@rti.org*; James Stephenson, RTI International; Carol Whisnant, RTI International; Diane Wagener, RTI International

Key Words: censored data, measurement error, partial correlation, proteomic expression, regression, typhoid vaccination

The association between immune response to typhoid vaccination and the expression levels of a set of defensins was investigated in this study. Defensins are a major class of antimicrobial peptides in mammals. The expression measurement of defensins was often heavily censored due to the detection and quantitation limits of the bioassay, and the accuracy of the immune response status was also limited by the accuracy of the bioassay. In addition, the data were not normally-distributed. To cope with these problems, the nonparametric test for first-order partial correlation with censored data (Akritas and Siebert, 1996) was first generalized to the higher order case and then applied. A censored regression model with defensin concentration as dependent variable and lognormal model errors was also applied for comparison. Simulation was conducted to evaluation the two approaches. The result is reported in this presentation.

Metabolic Profiling of Prostate Tissue Using the K-Means Cluster Analysis

✱Helena Gurascier, University of California, San Francisco, UCSF China Basin Campus, Radiology Department, 185 Berry Street, Suite 350, San Francisco, CA 94143-0946, *helena.gurascier@radiology.ucsf.edu*; Vickie Y. Zhang, University of California, San Francisco; Ying Lu, University of California, San Francisco; John Kurhanewicz, University of California, San Francisco

Key Words: k-Mean Cluster Analysis, Discriminant Factor Analysis, Metabolic Profiling, Diagnosis, Prostate Cancer

As PSA is not a specific cancer biomarker, it does not detect prostate cancer in 60%-75% of men and is unable to differentiate aggressive from non-aggressive cancers. This paper investigates metabolic profiling of prostate tissue simultaneously using 18 metabolite concentrations. 108 TRUS guided biopsies were obtained from human prostates and frozen to -80oC. 1D spectra were acquired on these samples using high resolution magic angle spinning (HR-MAS) spectroscopy. The concentrations were quantified using the algorithm QUEST. A k-Means Cluster Analysis was performed followed by a Discriminant Factor Analysis to obtain a 2-D plot. Three clusters were identified providing information on concentrations in the glandular, stromal and cancerous prostate tissue. These results indicate that prostate tissue has a strong biochemical signature that could provide a biomarker to detect cancer.

366 Bayesian Methods in Drug Development ●▲

Biopharmaceutical Section, Section on Bayesian Statistical Science, Section on Health Policy Statistics, Biometrics Section

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Clinical Trial Monitoring with Bayesian Hypothesis Testing

✱John D. Cook, M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Box 447, Houston, TX 77030, *cook@mdanderson.org*

Key Words: trial monitoring, Bayesian, hypothesis testing

Bayesian methods for sequential monitoring of clinical trials are often based on random inequalities. We present an alternative approach based on hypothesis testing with nonlocal priors. The test-based methods appear to have better operating characteristics than their inequality-based counterparts, based on asymptotic properties for large samples and simulation studies for small samples.

Bayesian Dose-Finding in Oncology for Drug Combinations by Copula Regression

✱Ying Yuan, M.D. Anderson Cancer Center, 1515 Holcombe Blvd, Unit 447, Dept. of Biostatistics, Houston, TX 77030, *yyuan@mdanderson.org*

Key Words: Adaptive design, Bayesian inference, Combining drugs, Continual reassessment

Treating patients with a combination of agents is becoming commonplace in cancer clinical trials, with biochemical synergism often the primary focus. In a typical drug combination trial, the toxicity profile of each individual drug has already been thoroughly studied in the single-agent trials, which naturally offers rich prior information. We propose a Bayesian adaptive design for dose-finding based on a copula-type model to account for the synergistic effect of two or more drugs in combination, while matching the toxicity profile of each single drug in the marginal model. To search for the maximum tolerated dose combination, we continuously update the posterior estimates for the toxicity probabilities of the combined doses. We conduct extensive simulation studies to examine the operating characteristics of the design.

Bayesian Tests for Synergy in Three-Agent Combination Chemotherapy Treatments

✱Melinda M. Holt, Sam Houston State University, Mathematics and Statistics, PO Box 2206, Huntsville, TX 77341, *mholt@shsu.edu*; Haunbiao Mo, Texas Woman's University

Key Words: Interaction, Fixed-dose Trials, Beta-binomial, Dose Response, Relative Potency

A current trend within in vitro chemotherapy studies is to combine two or more existing drugs in hopes that the mixture will be synergistic. We are specifically interested in determining whether or not three fruit enzymes act synergistically in combination; that is, whether or not the response to the combination is greater than expected under some chosen model of noninteraction. In order to identify synergy, one must establish a mathematical model for noninteraction with which to compare. Because these agents are believed to function via different direct mechanisms, the models of independence and effect summation are considered most appropriate. In addition, the agents of interest have been studied extensively alone and in pairs. Thus, Bayesian methods will be presented, including operational characteristics comparing results from the two models for several prior structures.

A Flexible Class of Models for Data Arising from a 'Thorough QT/QTc Study'

✱Suraj P. Anand, North Carolina State University, 2300 Avent Ferry Road, Apt N1, Raleigh, NC 27606, *suraj.anand@gmail.com*; Sujit Ghosh, North Carolina State University

Key Words: Bayesian Model Selection, DIC, ICH E14, MCMC, QT/QTc, thorough QT/QTc Study

The standard methods for analyzing data arising from a "thorough QT/QTc study" are based on multivariate normal models with common variance structure for both drug and placebo. Such modeling assumptions may be violated and when the sample sizes are small the statistical inference can be sensitive to such stringent assumptions. This article proposes a flexible class of parametric models to address the above mentioned limitations of the currently used models. A Bayesian methodology is used for data analysis and models are compared using the deviance information criteria (DIC). Superior performance of the proposed models over the current models is illustrated through a real data set obtained from a GlaxoSmithKline conducted thorough QT study.

Prior Estimation for Empirical Bayes Binomial Models

✱Boris G. Zaslavsky, U.S. Food and Drug Administration, 1401 Rockville Pike, H-219, Rockville, MD 20879, *Boris.zaslavsky@fda.hhs.gov*

Key Words: beta distribution, beta - binomial distribution, maximum likelihood, sample size

With the high cost of clinical trials, traditional methods of sample size determination may be too conservative. Bayesian methods are often used in the clinical trial environment to reduce required sample sizes and/or increase power. Bayesian methods are vital in trials where the observed results have to be extended on more extreme values which are unavailable for ethical reasons. The Beta distribution is a natural prior for binomial models. Under the empirical Bayes approach, the parameters of this distribution are the maximum likelihood estimator of the marginal beta-binomial distribution. For the sample size calculations, the maximum likelihood solution should be adjusted by a discount factor to reflect a partial exchangeability of historical trials as opposed to current studies. We suggest to measure the exchangeability.

367 Transportation, Visualization, Equity Premium Forecasting, IT, and Structural Breaks

Business and Economics Statistics Section, Section on Statistical Graphics

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

Estimation of Structural Breaks in Nonstationary Time Series: The Theory Behind AutoPARM

✱Stacey Hancock, Colorado State University, 1850 Laporte Ave Apt D22, Fort Collins, CO 80521, *staceyallayne@gmail.com*; Richard A. Davis, Columbia University; Yi-Ching Yao, Institute of Statistical Science, Academia Sinica

Key Words: time series, changepoint analysis

Many time series exhibit structural breaks, jumping to different stationary processes at times called “changepoints.” The goal of this research is to estimate where these structural breaks occur and to provide a model for the data within each stationary segment. The program AutoPARM (Automatic Piecewise AutoRegressive Modeling), developed by Davis, Lee, and Rodriguez-Yam (2006), uses the minimum description length principle to estimate the number and locations of changepoints in a time series by fitting autoregressive models to each segment. This research shows that when the true underlying model is segmented autoregressive, the estimates obtained by AutoPARM are consistent. Under a more general time series model exhibiting structural breaks, AutoPARM’s estimates demonstrate consistency-like properties.

Sequential Sampling for Pricing Consistency of Enterprise IT Solutions with Attached Services

✱Jerry Shan, Hewlett-Packard Company, 1501 Page Mill Road, Palo Alto, CA 94303, *jerry.shan@hp.com*

Key Words: sequential sampling, estimation, IT solution pricing

A customer IT solution that an enterprise provides can be a mix of hardware products (such as servers, storage, software, and PCs) and services. Consistency in pricing of these various elements is a key aspect to achieving certain strategic business objectives. This paper presents a statistical sampling based solution that we developed, with the objective of proving pricing consistency for one service deliverable. The solution was developed with a multi-stage, sequential sampling and estimation approach. In this paper, we will also report the encouraging test results with simulation data.

Communicating Quantitative Information Through Visualization

✱Krista L. Olson, BAE Systems, 8201 Greensboro Drive, McLean, VA 22102, *krista.olson@baesystems.com*

Key Words: Visualization, Graphs, Perception, Communication, Charts, PowerPoint

The most innovative analysis is of little use if it is not converted to information that can be effectively communicated. Visualization of information such as charts, graphs, and maps, if properly designed, can communicate complex information clearly and concisely. This paper pulls together the research done by renowned experts in the field of visualization such as Stephen Kosslyn, Colin Ware, Edward Tufte, William Cleveland, and others, into six recommended best practices for quantitative presentations. The author’s experience in using these techniques in the business world is included. Extensive examples are provided, encouraging interaction with the audience. The specific communication forum that is discussed is MS PowerPoint, one of the most used and criticized formats used today. This presentation will benefit anyone who ever finds themselves making presentations to others.

A Statistical Analysis of Ambulance Travel Times

✱Dawit Zerom, Mihaylo College of Business and Economics, LH 526, California State University Fullerton, P.O.Box-6848, Fullerton, CA 92834, *dzerom@fullerton.edu*

Key Words: Emergency Services, Conditional Distribution, Travel Times, Semi-parametric models

An important component of virtually any model of emergency service operations (an analytic model or optimization model) is the travel time of vehicles, especially of those en route to a call. In this paper, we present a semiparametric regression model for estimating travel times for emergency services using unique data from ambulance operations of the city of Calgary, Canada. The model allows nonlinearity of travel time-distance relationship, heterogeneity of variance, non-normality of the conditional distribution of travel times, the effect of other factors such as time of the day, daily trends, etc. Previous approaches of modeling travel times ignore the unique features of the conditional distribution of travel time, especially those of the conditional variance and excess kurtosis. Ignoring the latter aspects of travel time distribution is shown to lead to significant loss of accuracy.

Modeling of Strategic Planning Related to Land Use

✱Chandra Aleong, Delaware State University, 1200 DuPont Highway, Dover, 19901, *caleong@desu.edu*; J. Aleong, University of Vermont

Key Words: transportation, environment, meta- analysis

The main objectives of this study are to: (1) research the extent of the sprawl dilemma in Delaware and determine the primary problems associated with sprawl; (2) review integrated policies aimed at achieving a possible balance among land-use, transportation and environment; (3) understand benchmark processes used for integrated environmental decision making in Vermont and other key states; and, (4) develop a framework for integrated environmental policy making that can be adopted by Delaware. A meta-analysis of scientific research and policy papers will provide the information for the development of a conceptual framework of integrated environmental policy.

Estimation and Comparative Analysis of Departure Delays for Better Decisionmaking

✱Yufeng Tu, University of Maryland and Touro Univeriisity, *ytu@rhsmith.umd.edu*

Key Words: smoothing spline, delay distribution, dynamic prediction, variance reduction, downstream ripple effect, varying coefficient regression

The U.S. National Airspace System (NAS) is inherently highly stochastic. Yet, many existing decision support tools for air traffic flow management take a deterministic approach to problem solving. Modeling the randomness of the delays will provide a more accurate picture of the airspace traffic situation, improve the prediction of the airspace congestion and advance the level of decision making in aviation systems. In this study, we identify the seasonal trend and daily propagation pattern for flight delays and explore and evaluate several approaches to estimate the delays based on their ability to reduce variance and their predictive accuracy. Smoothing splines and varying-coefficient regressions are used in the study. Our models show significant improvement over the current model that the Enhanced Traffic Management System (ETMS) is currently using.

Out-of-Sample Equity Premium Prediction: Consistently Beating the Historical Average

✱Guofu Zhou, Washington University, Olin School of Business, Washington University, St. Louis, MO 63130, *zhou@wustl.edu*; David Rapach, Saint Louis University; Jack Strauss, Saint Louis University

Key Words: equity premium, predictability, Combination forecasts, Covariates

In this paper, we propose two approaches—combination forecasts and covariate estimation—to improve out-of-sample equity premium forecasts based on economic variables. These approaches accommodate structural instability and use additional information. We find that despite the failure of individual predictive regression model forecasts to outperform the historical average, combinations of individual model forecasts deliver statistically and economically significant out-of-sample gains relative to the historical average on a consistent basis over time. Forming combination forecasts from individual models estimated using covariates and with Campbell and Thompson (2007) restrictions imposed typically leads to further out-of-sample gains.

368 Invited Poster Presentations: Modeling in Ecology

Section on Statistics and the Environment

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

A Hierarchical Bayesian Modeling Approach for Analysis of Large River Fish Movement Data

✱Ali Arab, Georgetown University, Department of Mathematics, Georgetown University, 305 St. Mary's, Washington, DC 20057, *aa577@georgetown.edu*; Christopher Wikle, University of Missouri; Mark Wildhaber, U.S. Geological Survey; Aaron Delonay, U.S. Geological Survey

Key Words: Hierarchical Bayesian, physical/statistical model, tracking study, PDE, ecology, finite element method

Hierarchical Bayesian modeling approaches provide a flexible and effective tool for modeling problems related to habitat management in fisheries. Using this modeling framework, one is able to account for uncertainties in data and model parameters. Intensive management of large rivers such as impoundments, flow regulation, and channelization for purposes of navigation, flood control, and power generation result in dramatic physical changes to the river corridor, often eliminating many acres of habitat for native fish. One common method used to evaluate the impact of such alterations and modifications of a river basin on large river fishes is tracking movement through the use of telemetry. We develop a hierarchical Bayesian model characterized by the PDE-based dynamics based on their Galerkin finite element method (FEM) representations to analyze data obtained from such studies.

Data-Model Integration for Forest Dynamics

✱Jarrett J. Barber, University of Wyoming, 1000 E. University Ave, Department of Statistics, Dept. 3332, Laramie, WY 82071, *jbarber8@uwyo.edu*; Kiona Ogel, University of Wyoming; Darren E. Gemoets, University of Wyoming

Key Words: Bayesian hierarchical model, climate change, data-model integration, forest dynamics, scaling

Does predicting forests' responses to climate change require an investigation of individual or species behavior, or is a coarser and computationally less burdensome approach adequate? Our tree growth model allows environmental factors to drive trees' species-specific physiological processes constrained by species specific allometries such as the relationship of height to diameter. We use a Bayesian hierarchical model to integrate a large forest inventory database and extensive literature information with the model to obtain posterior estimates of model parameters. Initial results suggest that understanding within- and between-species variability in model parameters is important to predicting forest responses. These results suggest that future work on scaling between individual trees and climate model grid cells should account for the variability within and between species.

Modeling Marine Mammal Movement Data

✱Devin Johnson, National Marine Mammal Laboratory, Alaska Fisheries Science Center, 7600 Sand Point Way NE, Seattle, WA 98115, *devin.johnson@noaa.gov*; Carey Kuhn, National Marine Mammal Laboratory

Key Words: Animal telemetry, Integrated Ornstien-Uhlenbeck process, Northern fur seal, Changepoint model, Movement

We present a continuous-time movement model for the analysis of marine mammal telemetry data. A change-point model is considered for modeling changes is behavior which affect movement rates. A data set from a northern fur seal telemetry program is analyzed.

Hard Core or Soft Core: On the Characterization of Animal Space Use

✱Mevin B. Hooten, Utah State University, Logan, UT 84322, *mevin.hooten@usu.edu*; Ryan Wilson, Utah State University; John A. Shivik, Utah State University

Key Words: spatial point process, clustering, animal ecology

The delineation of animal home ranges is a long studied topic where kernel density based methods have been widely employed. Moreover, certain methods have been traditionally utilized for characterizing animal core areas, though a robust method, spanning species and environments, is still needed. We present a new method for statistical core area estimation when available data are limited to temporally independent animal locations with continuous spatial support on a bounded region. A key concept to this method is the optimal partitioning of a clustered spatial point process into a set of two complete spatially random point processes such that the union of their support is equal to the home range. Additionally, we account for uncertainty in the core area delineation by allowing it to be a random set that is parameterized and estimated with Bayesian methods.

Hierarchical Models for the Estimation of Manatee Abundance from Aerial Surveys

✱Robert M. Dorazio, U.S. Geological Survey/University of Florida, Department of Statistics, University of Florida, Gainesville, FL 32611-0339, *bdorazio@ufl.edu*

Key Words: Abundance, Aerial survey, Double-observer sampling, Habitat, Manatee, Removal sampling

Predictions of manatee abundance as a function of habitat characteristics are needed for making conservation decisions for this threatened species. The spatial distribution of manatees in southwest Florida varies seasonally in response to changes in salinity, water temperature, food availability, and other factors related to habitat. Consequently, aerial surveys were developed to estimate abundance of manatees in spatially referenced sample units using a combination of sampling protocols. Groups of manatees were detected using double-observers, and the number of manatees in each group were detected by repeated circling to yield a sequence of "removal" counts. Thus, both kinds of counts (i.e., those of groups and those of manatees within groups) were spatially referenced.

Bayesian Hierarchical Models for the Lower Trophic Ecosystem in the North Pacific Ocean

✱Christopher Wikle, University of Missouri, MO 65211, *wiklec@missouri.edu*; Ralph Milliff, Northwest Research Associates, Inc.

Key Words: ocean, ecosystem, dynamical, nonlinear, hierarchical

A wide variety of physical-biological models of varying complexity have been developed for components of the U.S. West Coast upwelling ecosystem. Such a physical-biological interface is effective for demonstrating high-resolution properties of phytoplankton bloom and zooplankton population dynamical response. We develop a simple (nitrogen, phytoplankton, zooplankton) stochastic model in a hierarchical Bayesian framework. The underlying system is multivariate and highly nonlinear. The model is applied to the upwelling region of the North Pacific and demonstrates the interplay between model formulation, data, and interpretation of posterior distributions in relatively simple stochastic models for a complicated physical-biological system. Ultimately, such models can be used for pan-regional syntheses and climate change impact studies for coastal ocean ecosystems.

369 Contributed Poster Presentations

Biometrics Section, IMS, Section on Nonparametric Statistics, Section on Physical and Engineering Sciences, Section on Quality and Productivity, Section on Statistical Computing, Section on Statistical Education

Wednesday, August 6, 8:30 a.m.–10:20 a.m.

A Step-Down Pairwise Comparison Procedure Using Medians and Permutation Tests

✱Scott Richter, The University of North Carolina at Greensboro, *sjricht2@uncg.edu*; Melinda McCann, Oklahoma State University

Key Words: permutation test, multiple comparisons

In this paper is developed a step-down procedure based on the single step median-based procedure of Richter & McCann (2007). The sampling distribution is based on the permutation distribution of the maximum median difference among all pairs. The step-down procedure provides greater power for detecting group location differences while still maintaining strong FWE control. The new procedure can also be more powerful than popular step-down procedures for comparing normal means, especially for heavy tailed distributions.

Using the Bootstrap To Select Active Factors in Unreplicated Factorial Experiments

✱Gerald Shaughnessy, University of Dayton, Dept of Mathematics, University of Dayton, Dayton, OH 45449-2316, *Gerald.shaughnessy@notes.udayton.edu*; Maher Qumsiyeh, Bethlehem University

Key Words: Bootstrap, Factorial Experiments

The bootstrap allows for confidence intervals to be constructed for the effect size of the active factors What was noise in the original experiment definitely stayed as noise This suggests that the bootstrap can be used to select the active factors, especially in doubtful situations.

Using a Mark-Recapture Experiment To Illustrate Methods of Point and Interval Estimation

✱Carlo Cosenza, California State University, East Bay, 19983 Santa Maria Ave, Apt #301, Castro Valley, CA 94546-4242, *crc27@cornell.edu*; Joshua Kerr, California State University, East Bay; Bruce Trumbo, California State University, East Bay

Key Words: Estimation of population size, capture-recapture, maximum likelihood estimation, bootstrap confidence interval, Bayesian interval estimation, R, elementary uses

Estimates of population size in a classical mark-recapture experiment are obtained by method of moments and maximum likelihood estimation. Traditional, bootstrap, and Bayesian methods of interval estimation are also illustrated. Methods are appropriate for advanced undergraduates and first-year MS students. R is used to do the intensive computation and to make figures.

An Algorithm for Unconstrained Quadratically Penalized Convex Optimization

✱Steven P. Ellis, Columbia University, Unit 42, NYSPI, 1051 Riverside Dr., New York City, NY 10032, *spe4@columbia.edu*

Key Words: optimization, machine learning, nonparametric estimation, statistical computing

A common way to approach (supervised) nonparametric estimation problems in machine learning is as follows. Bound the empirical risk or likelihood above by a convex functional, L, on a reproducing kernel Hilbert space, H, and (*) minimize over H the sum of L and a penalty proportional to the squared norm on H. (*) is a convex minimization problem. The theory of convex optimization is well developed. However, convex optimization software tends to be specialized and hard to use. We describe an algorithm ("QQMM") designed to solve unconstrained optimization problems of the form (*), particularly when computing L is expensive. QQMM does not require the Hessian matrix of L, but does use (sub)gradients. A common, general purpose optimization method is BFGS. QQMM is compared to BFGS and their relationship discussed. Examples are presented.

Numerical Error in ODEs

✱Clyde Martin, Texas Tech University, Dept. of Mathematics and Statistics, Lubbock, TX 79409, *clyde.f.martin@ttu.edu*; Bo He, Texas Tech University

Key Words: differential equations, numerical error, switching

An autoregressive integrated moving average model (ARIMA) is developed to study the statistical behavior of the numerical error generated from three fourth-order ODE solvers: Milne's method, Adams-Bashforth method and a new method which randomly switches between Milne and Adams-Bashforth methods. With the error data based on four differential equations, we identify an ARIMA model for each data series. Results show that some of data series can be well described by ARIMA models but others can not. Based on the mathematical form of the numerical error, other statistical models should be investigated. Finally we assess the multivariate normality of the sample mean vectors which are generated by the switching method as an application of the multivariate central limit theorem. It is shown that the error for the switching method is well behaved.

Differences in Computations of Sample Quantiles

✱Michael Joner, The Procter & Gamble Company, 8700 S Mason-Montgomery Rd, Mason, OH 45040, *joner.md@pg.com*; Eugene Shine, Savannah River National Laboratory

Key Words: quantile, software, comparison

Sample quantiles can be found in at least 9 ways using one or two order statistics (Hyndman and Fan, 1996). Different methods are in use by various software packages. This can be a source of confusion when multiple packages are in use because the methods often produce different estimates. We review some of the technical issues in sample quantile estimation.

Asymptotic Expansions in Mean and Covariance Structure Analysis

✱Haruhiko Ogasawara, Otaru University of Commerce, 3-5-21 Midori, Otaru, International 047-8501 Japan, *hogasa@res.otaru-uc.ac.jp*

Key Words: mean and covariance structure, factor means, Edgeworth expansion, Cornish-Fisher expansion, bias, skewness and kurtosis

Asymptotic expansions of the distributions of parameter estimators in mean and covariance structures are derived. The parameters may be common to or specific in means and covariances of observable variables. The means are possibly structured by the common/specific parameters. First, the distributions of the parameter estimators standardized by the population asymptotic standard errors are expanded using the single- and the two-term Edgeworth expansions. In practice, the pivotal statistic or the Studentized estimator with the asymptotically distribution free standard error is of interest. An asymptotic distribution of the pivotal statistic is also derived by the Cornish-Fisher expansion. Simulations are performed for a factor analysis model with nonzero factor means to see the accuracy of the asymptotic expansions in finite samples.

A Comparison of Two Boxplot Methods for Detecting Outliers That Adjust for Sample Size and, When Appropriate, for Asymmetry

✱Nancy J. Carter, California State University, Chico, Department of Mathematics & Statistics, 400 W. First Street, Chico, CA 95929-0525, *ncarter@csuchico.edu*; Neil C. Schwertman, California State University, Chico; Terry L. Kiser, California State University, Chico

Key Words: boxplot, masking, swamping, Monte Carlo

It is important to identify outliers. One of the easiest and most useful methods is based on the boxplot. Two methods, one by Carling and another by Schwertman and de Silva, adjust the boxplot method for sample size and skewness. In this paper, the two procedures are compared both by theoretical computations and Monte Carlo simulations. Carling's method is to the Schwertman and de Silva procedure as comparisonwise versus experimentwise error rate is for multiple comparisons. The Carling approach is superior in avoiding masking outliers while the Schwertman and de Silva procedure is much better at reducing swamping.

Parallel Processing in Artificial Intelligence

✱Morteza Marzjarani, Saginaw Valley State University, SE-179 Computer Science Department, 7400 Bay Road, University Center, MI 48710, *marzjara@svsu.edu*; Joshua Urbain, Saginaw Valley State University; Joshua Cieszlak, Saginaw Valley State University

Key Words: Expert Systems, Computers

By researching and creating a way to make statistical methods into distributed jobs, we can distribute work among many cost-effective computers, rather than forcing users to wait long periods of time or use expensive supercomputers for statistic applications. Computing power for a consumer in this day and age can be bought for low prices, but larger-scale operations can be quite expensive. By using distributed computing, one is able to separate pieces of a program onto multiple computers. Statistic operations can be very time-consuming and power-hungry, making it a great match for distributed computing. The methods created by this research can then be used by anyone in the many fields that use complex applications: computer architecture, databases, biology, arts, English, etc.

Entropy Estimation of Multimodal Circular Distributions

✱Shengqiao Li, Centers for Disease Control and Prevention, Morgantown, WV 26505, *SWL4@CDC.GOV*; Robert Mnatsakanov, Centers for Disease Control and Prevention; Adam Fedorowicz, Centers for Disease Control and Prevention; Michael E. Andrew, National Institute for Occupational Safety and Health

Key Words: Entropy, Circular Distribution, Generalized von Mises Distribution, EM Algorithm, KNN, Spacings

In the entropy estimation problem of multimodal circular distributions, four methods are proposed. They are based on: 1) generalized von Mises (GvM) model; 2) finite mixtures of von Mises (MvM) distributions; 3) circular spacings of order k (KCS), and 4) k-nearest neighbor (KNN) construction. GvM can be skewed and/or multimodal. Its parameters are estimated numerically via MLE approach. In the case of MvM, an EM algorithm is derived and applied to estimate the model parameters. These four methods are compared through simulations. It is shown that the two parametric models and KCS method perform quite well and they are a little bit better than KNN estimator in terms of RMSE. KCS approach is very simple and fast; therefore it is recommended in our study. Finally, the rotational entropies of four different dihedral angles of (S, S)-tartaric acid molecule are estimated using these methods.

Exact Bounds on the Probability That One Variable Exceeds Another

✱Robert Lew, VA Boston Healthcare System, 150 S. Huntington Ave., Boston, MA 02130, *robert.lew2@va.gov*

Key Words: exact bounds, inequalities, Tchebycheff systems, Wilcoxon test, unimodality

In rank tests such as the Wilcoxon test, the expression $P(W < = Y)$arises, where W and Y have unit variance normal distributions: $W \sim N(m,1)$, $Y \sim N(0,1)$, and $m > 0$. To assess robustness we find exact bounds under two sets of assumptions; Set1)$W \sim N(0,1)$ and Y has unit variance and is unimodal and symmetric at zero and Set2) W and Y have unit variance and W-m and Y are unimodal and symmetric at zero. Under Set1 Tchebycheff system results yield exact algebraic bounds. Under Set2 a more general result yields complicated exact algebraic bounds. A simple search avoids the algebra. The method for Set2 applies to nearly any expression in W and Y so can be used whenever numerical results suffice.

Constrained Maximum Likelihood Estimation for Logistic Regression

Martin Levy, Universtiy of Cincinnati; ✱ZhiYuan Dong, Universtiy of Cincinnati, Department of Quantitative Analysis, PO Box 210130, Cincinnati, OH 45221, *dongzn@email.uc.edu*; Yan Yu, Universtiy of Cincinnati; Tim Keyes, GE Capitol

Key Words: Linear Programming, Optimization

Often prior information on event probabilities is available to logistic regression analysts. Classical maximum likelihood estimation ignores this information while Bayesian approaches can take full advantage of it. An intermediate approach couches the information in terms of constraints that either order the probabilities or set bounds on them. When these constraints translate into linear constraints on the estimated regression parameters, one can take computational advantage of common statistical packages to estimate them. Based on simulations we find that even with modest sample size and few constraints estimation accuracy is greatly improved over unconstrained MLEs based on a simple metric. Moreover, nice asymptotic properties of the constrained MLEs are retained. With bounding constraints, Bayesians can construe the bounds as noninformative priors on the event probabilities.

Determining Minimum Sample Sizes To Achieve Central Limit Theorem Closeness When Sampling from Various Populations

Jamis Perrett, Texas A&M University; ✱Steve Hoff, University of Northern Colorado, CB 124, University of Northern Colorado, Greeley, CO 80639, *hoff2584@blue.unco.edu*

Key Words: Normality Assumption, Minimum Sample Size, Central Limit Theorem

One version of the Central Limit Theorem states that when repeatedly sampling independently from any distribution with finite first and second moments, and for some large sample size (n), the sample mean possesses an approximate Gaussian distribution. How close is "approximate" and how large is "large"? Using techniques from mathematical statistics and from simulation sampling, the researchers have developed a technique to assess minimum sample sizes which are sufficient to insure approximate closeness for selected populations both symmetric and skewed, with domains both finite and infinite, and with both light and heavy tails. Simulation routines are designed in a manner consistent with minimum sample sizes determined from theoretical results.

Classification of Closed Contours

✱Fu-Chih Cheng, North Dakota State University, 203B WALDRON HALL, FARGO, ND 58105, *fuchih.cheng@ndsu.edu*

Key Words: Classification, Contour, Wavelets, Fourier

One of the important aspects in pattern recognition is shape classification. Given objects or contours, one would like to locate their corresponding memberships. Traditionally, one would approach this issue by taking the measurements from objects of interest or by taking the coefficients of the shape representation, namely coefficients of Fourier descriptors. Once the measurements of contours or coefficients of Fourier descriptor of contours are obtained, statistical classification procedures can be used. In our research, wavelet-based contour classification is introduced. More specifically, the wavelet transformation is applied onto the contour, and the corresponding coefficients are used for the purpose of statistical classification. The simulation study shows that the wavelet-based method is superior to the other two methods.

An Artificial Immune Network–Based Classification Approach to ECG Monitoring Applications

Honggang Wang, Univeristy of Nebraska-Lincoln; ✱Hua Fang, University of Nebraska-Lincoln

Key Words: Artificial Immune Network, Classification, ECG Pattern Recognition

Electrocardiograph (ECG) is a widely used tool for cardiac monitoring. As traditional ECG curves only represent a short time sampling of patient data, irregular and intermittent cardiac conditions may not be identifiable and therefore real time monitoring is necessary. However, tremendous ECG data are generated by the real time monitoring and immediate health care are needed for the patients, which pose a substantial computational challenge for ECG pattern recognition. Artificial Immune Network (AIN) classifier is proposed to automatically discern and reduce the ECG data size using MemeoryCell representation. This approach combines PCA and K Nearest Neighbor to conduct projection pursuit and supervised classification based on the Memory Cells. Following a real patient study, a simulation study is conducted to exhibit the advantages of our approach compared to traditional methods.

Using Semiparametric Varying Coefficient Models To Investigate Interactions of Toxic Exposure and Nutritional Covariates in a Study of Neurodevelopmental Outcomes

✱Miranda L. Lynch, University of Rochester, Department of Biostatistics and Computational Biology, 601 Elmwood Avenue, Box 630, Rochester, NY 14642, *miranda_lynch@urmc.rochester.edu*; Li-Shan Huang, University of Rochester

Key Words: Varying coefficient models, Semiparametric models, Neurodevelopmental outcomes

Varying coefficient models are a flexible method for quantifying interactions by allowing the linear association between a primary regressor and a response to be modeled via linear model coefficients that are themselves smooth functions of a separate effect-modifying covariate. The Seychelles Child Development Study (SCDS) is a long-term study of the effects of prenatal mercury (Hg) exposure on measures of neurodevelopmental endpoints in children. Recent work in the SCDS has suggested that deleterious effects of Hg on development may be attenuated by complicated interactions of Hg with one or more nutrition covariates. Interpretation of these interactions is crucial to understanding the role of nutrition in modifying the effects of Hg on developmental outcomes. These interactions will be explored using semiparametric varying coefficient models based on penalized splines.

Performance of Some Statistical Methods for Testing Equality of Parameters in Zero-Inflated Continuous Distributions

✱Lei Zhang, Mississippi State Department of Health, Office of Health Data & Research, 570 E. Woodrow Wilson, PO Box 1700, Jackson, MS 39215-1700, *lei.zhang@msdh.state.ms.us*; William D. Johnson, Lousiana State University System; Jixiang Wu, Mississippi State University

Key Words: Continuous, Distribution, Zero-Inflated, Statistical, Methods, Equality

We evaluated the properties of five statistical methods for testing equality among populations with zero-inflated continuous distributions. These tests are based on likelihood ratio (LR), Wald, parametric jackknife (PJ), nonparametric jackknife (NPJ), and central limit theorem (CLT) statistics. We investigated their statistical properties using Monte Carlo simulated data from distributions with gamma, exponential, and log-normal density functions. These five statistical tests are compared in terms of the empirical Type I errors and powers estimated through 10,000 repeated simulated samples for carefully selected configurations of parameters. The LR test performed the best since its empirical Type I errors were close to the preset nominal 0.05 level and demonstrated desirable power for rejecting alternative hypotheses.

Semiparametric Inference of Linear Transformation Models with Length-Biased Censored Data

✱Jane Paik, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, *jane@stat.columbia.edu*; Zhiliang Ying, Columbia University

Key Words: Semiparametric linear transformation model, biased sampling, censoring

We propose estimating equations for the simultaneous estimation of regression parameters and the transformation function in semiparametric linear transformation models when dealing with biased sampling in censored data. Existing estimation procedures for censored data using linear transformation models yield biased estimators for regression parameters of interest. Our approach is motivated by the unified estimation procedure proposed by Chen et al. (2002). The proposed estimators for the regression parameters are proven to be consistent and asymptotically normal. The variance-covariance matrix has a closed form which can be consistently estimated. The finite sample performances under various scenarios are assessed through simulation studies, which indicate that the proposed estimators give negligible bias and correct coverage probabilities. The method is also applied to a real data set.

A Distribution-Free, One-Sample Test for Equivalence

✱Jesse Frey, Villanova University, Mathematical Sciences Department, 800 Lancaster Ave, Villanova, PA 19085, *jesse.frey@villanova.edu*

Key Words: bias, confidence bands, goodness-of-fit, Kolmogorov-Smirnov test

There are many exact distribution-free goodness-of-fit tests, but no equivalence testing analogues. This paper fills this gap by developing an exact one-sample distribution-free equivalence test for use with continuous distributions. We consider two continuous distributions equivalent if the pointwise distances between their distribution functions never exceed some specified constant, and we test equivalence using the sup of the pointwise distances between the empirical distribution function and the fully specified distribution of interest. The new test is much more powerful than the naive distribution-free equivalence test obtained from the Kolmogorov-Smirnov confidence band for the distribution function, and inversion of the test leads to distribution-free confidence bands for the unknown distribution function that are centered at the fully-specified distribution of interest.

A Nonparametric Version of the Multivariate Likelihood Ratio Test (Wilks Lambda)

✱Chunxu Liu, University of Kentucky, Department of Statistics, 814 Patterson Office Tower, Lexington, KY 40506-0027, *chunxu@ms.uky.edu*; Arne Bathke, University of Kentucky

Key Words: Rank Test, Multivariate Statistics, Likelihood Ratio, Wilks Lambda, Asymptotics, Small Sample Approximations

We propose a nonparametric version of the multivariate likelihood ratio test and investigate its asymptotic properties. We also evaluate different small sample approximations by means of simulations whose settings are motivated by clinical trials and agricultural data. Application is illustrated by an example.

The Rank Transforms and Tests of Interaction for Repeated Measures Experiments with Various Covariance Structures

✱Jennifer Bryan, Oklahoma Christian University, 1003 Dustin Drive, Edmond, OK 73003, *jennifer.bryan@oc.edu*; Mark E. Payton, Oklahoma State University

Key Words: Repeated measures, Aligned Rank Transform, Nonparametric methods

The covariance structure of a repeated measures design can be simple or very complicated. In analyzing repeated measures, rank transformations can be an alternative to the standard tests performed on the raw data. An alternative to utilizing the common rank transform when testing for interaction is the aligned rank procedure in which the estimate for the interaction effect is adjusted for the observed main effects. The question arises as to how the covariance structure may affect the aligned rank transform procedure when analyzing repeated measures, specifically the test of interaction. Simulation results comparing test sizes and power of the rank transform, aligned rank and raw data analyses will be presented.

Testing the Structure of a Covariance Matrix

✱Tracy Morris, University of Central Oklahoma, *tmorris2@ucok.edu*; Mark E. Payton, Oklahoma State University

Key Words: permutation test, covariance matrix, nonparametric test, randomization test

The classical method for testing the structure of a covariance matrix involves a likelihood ratio test (LRT). These tests have many limitations, including the need for large sample sizes and the requirement of a random sample from a multivariate normal population. The LRT is also undefined for cases in which the sample size is not greater than the number of variables. There are many situations in which at least one of these conditions is violated. In such situations, permutation tests are appropriate. This research involves the development of a permutation test for the structure of a covariance matrix. Samples of various sizes, numbers of variables, and underlying covariance structures were simulated from multiple distributions. In each case, the type I error rates and power were examined.

Nonparametric Estimation of Copula Density Functions

✱Wenhao Gui, Florida State University, Department of Statistics, Tallahassee, FL 32306, *wenhao@stat.fsu.edu*

Key Words: Copula, Orthogonal Series, Kernel Estimator, Nonparametric

Copulas are the functions linking univariate marginals to their joint distribution functions. They measure the dependence among components of random vectors and become a popular tool in multivariate modeling. In this paper, we propose a new nonparametric method to estimate copula density functions. The proposed estimators are based on orthogonal series. Under mild conditions, the asymptotic properties of estimators are given. We also consider hard thresholding and soft thresholding for sparse representations. We compare the estimator with the kernel estimator. Simulations and applications to real data are given.

Ordered Inference Using Observed Confidence Levels

✱Alan M. Polansky, Northern Illinois University, Division of Statistics, De Kalb, IL 60115, *polansky@math.niu.edu*

Key Words: Bootstrap, Edgeworth Expansion, Normal Model, Restricted Inference, Multiple Comparisons, Error Rate

Statistical inference on the ordering of the elements of a mean vector is an important issue in many applied problems. Many statistical tests of hypotheses to detect these orderings have been developed both within the normal model, and outside the normal model using nonparametric methods. Estimates as well as confidence regions have also been developed for the mean vector under constraints imposed by these ordering models. In order to attempt to distinguish between ordered models, multiple testing procedures are often used to control the overall error rate of the sequence of tests. This paper shows how observed confidence levels allow for the exploration of very general models for the ordering of the elements of a mean vector without the need for specialized asymptotic theory or multiple testing methods. The methods are applied to several well-known examples.

Adaptive Quasi-Monte Carlo Sampling: Construction and Applications

✱Ben Haaland, University of Wisconsin-Madison, Madison, WI 53719, *haaland@stat.wisc.edu*; Zhiguang (Peter) Qian, University of Wisconsin-Madison

Key Words: quasi-monte carlo, space filling designs, sequential designs, nested designs, computer experiments, nets

Sequences of nested, extendable space filling designs are useful in statistics, computer science, and numerical analysis. Results for nesting and extending a class of designs with good uniformity properties, the (t,m,s)-nets and (t,s)-sequences, are presented. Applications and examples in numerical integration, design and modeling for computer experiments with multiple levels of accuracy and expense, and stochastic programming are given.

Separating Borrowing Information and Forming Contrasts: Nonparametric Inference for Arbitrary Functionals of Survival

✱Kyle Rudser, University of Washington, Department of Biostatistics, Campus Mail Stop 357232, Seattle, WA 98195, *rudserk@u.washington.edu*; Scott S. Emerson, University of Washington; M. L. LeBlanc, Fred Hutchinson Cancer Research Center

Key Words: Semiparametric and nonparametric methods

The hazard ratio is commonly used for comparing survival distributions across groups. While easily estimated in the presence of censored data, it does not allow for the clinical relevance of differences in survival across groups to be easily judged. We consider an approach to nonparametric inference for clinically meaningful functionals of a survivor distribution (e.g., restricted mean, quantiles). In this approach we use different models to borrow information across sparse data than to form contrasts. Linear contrasts are evaluated and compared on root mean squared error and coverage between approaches using nonparametric recursive partitioning, Cox's proportional hazards, and Buckley-James' linear regression with censored data. The nonparametric approach was superior when semiparametric model assumptions were violated, and had a slight loss of efficiency when such assumptions do hold.

370 The Analysis of Wikipedia Data ●

Section on Statistical Computing
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Wikipedia as a Model for ASA Journals

✱David Banks, Duke University, Dept. of Stat. Science, Box 90251, Durham, NC 27708, *banks@stat.duke.edu*

Key Words: electronic publication, growth models

Wikipedia provides a successful model for volunteer creation of collaborative content. Another example is R, now almost universally used in statistics graduate programs. This talk discusses the implications of these models for how the ASA can move further into the electronic age, providing new services and functionality to its members and to allied scientists.

Quality Assurance in Information Submissions: The Case of Wikipedia

✱Robert Warren, University of Waterloo, Waterloo, ON N2L3G1 Canada, *rhwarren+jsm@uwaterloo.ca*

Key Words: Social Network Analysis, Quality Control, Wikipedia, Performance Metrics

Wikipedia makes use of a distributed editing system to ensure quality within the published articles. This talk is an performance analysis of the quality control processes built into Wikipedia, with special attention paid to the antagonist behavior observed of some one-off editors.

Can Statistical Learning Sustain Wikipedia's Model of Growth?

✱Edo Airoldi, Princeton University, Carl Icahn Laboratory, Princeton, NJ 08544, *eairoldi@princeton.edu*

Key Words: Network structure, Semantic structure, Evolution, Latent variables, Data integration, Semi-supervised learning

The administration and daily use of large online databases of textual information require a substantial effort to organize content for facilitating browsing (*www.jstor.org*) or categorization of new material (*www.wikipedia.org*). Statistical learning tools that infer semantic patterns from text hold the promise of automating, at least in part, organization and re-organization of content. This talk discusses methodological issues at the core of such tools, and strategies for assessing their utility in sustaining automation, in the context of a case study.

371 The State of Statistics Education in Schools (K–12) Around the World ●▲

International Association for Statistical Education,
Section on Statistical Education
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

GAISEing into the Future in New Zealand

✱Chris J. Wild, University of Auckland, Department of Statistics, 38 Princes Street, Auckland, International 1010 New Zealand, *wild@stat.auckland.ac.nz*; Nicholas J. Horton, Smith College

Key Words: informal inference, GAISE, school curriculum, supporting materials, statistical literacy, teacher development

The 2005 GAISE report laid out a vision for Pre-K–12 US statistics education. Subsequently, New Zealand began developing a detailed national curriculum in the spirit of GAISE with increased emphasis on statistics. A high-level curriculum specification was finalized in 2007, and the next level of detail is under development. Translating high-level, forward-looking principles into a detailed reform curriculum is an enormous challenge. We need to plant the seeds in earlier years that can germinate and flower into fully fledged concepts at later years. Despite the NZ curriculum sharing a very similar philosophical base to GAISE, there are intriguing points of departure. We will compare and contrast GAISE and developments in NZ; confront the difficulties with teacher preparation, professional development, and creation of materials; and identify areas where we can leverage off one another.

Efforts of Government Statistical Agencies Across the World to Complement School Curricula

✱Mary Townsend, Statistics Canada, 1 B10 Jean Talon Building, 170 Tunney Pasture Parkway, Ottawa, Ontario, K1A 0T6 Canada, *Mary.Townsend@statcan.ca*

Key Words: Statistical Agencies, Government, data handling, chance, student's data, media graphs

Government statistical agencies in many countries (alone or in coordination with academic institutions) have actively engaged in helping school teachers and students gain confidence in statistical literacy. Contrary to myth, these efforts have not only consisted of teaching what the inflation rate, the Gini coefficient, or other indicators are, but on active data handling and chance activities based on students' data. In some cases, interpretation of the use of statistics in the media, graphical interpretations of information, and progressive pedagogical contents have been used. This talk presents the results of a survey of these efforts and their effectiveness in complementing or encouraging school curricula on data and chance.

How Can IASE Have an Impact on Statistics Education in Schools?

✱Larry Weldon, Simon Fraser University, Dept of Statistics and Actuarial Science, Burbany, V5A 1S6, British Columbia, Canada, *weldon@sfu.ca*

Key Words: K-12, teaching statistics, survey, impact analysis, assessment, measurement

The ICOTS conferences and ISI/IASE Satellite meetings provide valuable printed resources on modern trends in statistics education. Activists in the schools can make use of these to modernize and improve their teaching, but the demands of school curricula are constraining. Both curricula and teachers of school statistics will not reform soon without reform at the tertiary level. But there are constraints to tertiary reform as well: textbook publisher conservatism and the priority of research in tertiary career progress. In this paper, we propose a way to overcome the inertia in tertiary curriculum reform: include more guided experiential learning and less technique-oriented lecturing, more projects and extended applications, and less text-book coverage. To facilitate transfer to the schools, IASE needs to sell this tertiary reform to the school activists.

372 Noether Award Invited Session

Noether Award Committee, Section on
Nonparametric Statistics
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

A New Approach to R-Estimation

✱Davy Paindaveine, Université Libre de Bruxelles, Brussels, B-1050 Belgium, *dpaindav@ulb.ac.be*; Marc Hallin, Université Libre de Bruxelles

Key Words: R-Estimators, Local asymptotic optimality, Robustness, Semiparametric statistics

Traditional R-estimators, defined via the minimization of rank-based objective functions, are unsatisfactory not only from a practical point of view (the implementation is numerically infeasible for high-dimensional parameters), but also from a theoretical point of view (root-n consistency is nontrivial). We suggest a rank-based adaptation of Le Cam's one-step construction of locally asymptotically optimal estimators. The new estimators are asymptotically equivalent to a random vector measurable with respect to the ranks associated with the true value of the parameter, and hence can be considered as genuine R-estimators. They are easy to compute and exhibit good finite sample properties. As expected, they are locally and asymptotically optimal under correctly specified densities. Finally, their Gaussian-score versions often uniformly dominate their parametric Gaussian counterparts.

Conditional U-Statistics with Applications in Discriminant Analysis, ARMA Processes, and Hidden Markov Models

✱Madan Puri, Indiana University, Rawles Hall, Department of Mathematics, Bloomington, IN 47405, *puri@indiana.edu*

Key Words: Nonparametrics, U-Statistics, ARMA, Discrimant Analysis, Hidden Markov Models

In memory of Gottfried Noether, a prominent scholar in nonparametric statistics, the Noether family and ASA established the Noether Awards in 2000. The Noether Awards are given to recognize and encourage scholarship in nonparametric statistics. This Noether Award invited session will be a session in which Madan Puri, the recipient of the Senior Noether Award, will present a major paper on nonparametric statistics.

373 The Role of Statisticians in Understanding Climate Change ●▲

Section on Statistics and the Environment, Scientific and Public Affairs Advisory Committee, WNAR, *CHANCE*

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

The ASA-Sponsored Workshop on Climate Change

✱David A. Marker, Westat, 1650 Research Blvd., Rockville, MD 20850, *DavidMarker@Westat.com*

Key Words: global warming, data quality, modeling

In October 2007, the ASA sponsored a workshop bringing together statisticians and atmospheric scientists to reach a consensus on the statistical issues in climate change research and to identify areas where statisticians could play a strong role in future research. The models on which climate change projections are based are complex and rely on data that vary greatly in quality. This presentation will summarize the topics discussed at the workshop and the recommendations that were developed for future research.

Democracy vs. Aristocracy in Model Projections for Climate Change

✱Claudia Tebaldi, National Center for Atmospheric Research, *tebaldi@ucar.edu*

We present work that has explored the correlation between several statistical summaries of model output for current climate and the models' future behavior in terms of amplitude of future trends and their variability. We also use a perfect model approach to try to get at the same issue—how current climate simulations can tell us something about future simulations' reliability. With this work, we hope to attack the crucial issues of merging models' projections in an informed fashion, while we wait for a time when all models will perform equally well and we will be satisfied with a democratic handling of each model's voting rights.

Where Are Statisticians in the Earth System?

✱Doug Nychka, National Center of Atmospheric Research, 1850 Table Mesa, Boulder, CO 80307-3000, *nychka@ucar.edu*

The past ten years have been a watershed in the geosciences where separate models for the ocean, atmosphere, land and ice, have been combined to understand the complex interactions of the Earth's climate system. Added to this physical perspective are biological, chemical and social dimensions, important for studying climate change due to human activities. This talk will give an overview of how statistical science can fit into this broad scientific effort and of the potential for new kinds of interdisciplinary collaborations. A primary contribution of statistics is in quantifying uncertainty throughout the modeling process and in comparing model predictions to sparse and irregular climate data. A intriguing new direction, however, is to introduce stochastic elements into the geophysical models themselves.

374 Measuring Health Care Disparities ●

Social Statistics Section, Section on Government Statistics, WNAR, Section on Health Policy Statistics, Section on Statistics in Epidemiology

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Understanding Disparities Within and Among Geographical and Health Care Units

✱Alan Zaslavsky, Harvard University Medical School, 180 Longwood Avenue, Dept of Health Care Policy, Boston, MA 02115-5821, *zaslavsky@hcp.med.harvard.edu*

Key Words: health care, disparities, geography, multilevel models, race

Health and health care disparities can originate at various levels within the social and healthcare systems. Variations among geographical areas and among health care units (health plans, hospitals, doctors) within the same geographical area may contribute to disparities if assignment to these units is associated with race/ethnicity. We consider analytic methods for identifying the contributions to disparities from these various levels, with examples from major health care quality measurement systems. We also discuss the interpretation of racial/ethnic disparities that are "mediated" through socioeconomic differences. Policy implications will be drawn out.

Application of the Peters-Belson Method for Estimating Disparities

✱Barry I. Graubard, National Cancer Institute, 6120 Executive Blvd, Room 8024, Bethesda, MD 20892, *graubarb@mail.nih.gov*; Sowmya R. Rao, MGH Biostatistics Center/The Institute for Health Policy; Joseph Gastwirth, The George Washington University

Key Words: health disparities, regression, complex samples, cancer screening, cancer incidence

The Peters-Belson (PB) method is used to estimate wage discrimination and disparities in cancer screening between an advantaged group (e.g., Caucasians) and a disadvantaged group (e.g., Hispanics). The PB method uses regression to estimate the proportion of the observed disparity that is not explained by the covariates used in the regression. This method first fits a regression model with individual-level covariates to the majority group and then uses the fitted model to estimate the expected values of the minority-group members had they been members of the majority group. We extend the PB method to using a sample with measured covariates from an external population. We illustrate the methods with cancer screening and cancer incidence data from the National Health Interview Survey and the National Health and Nutrition Examination Survey that have weighted multistage cluster samples.

Improved Analysis of Weight-Loss Interventions for African-American Women

✱Justine Shults, University of Pennsylvania, 610 Blockley Hall, School of Medicine, 423 Guardian Drive, Philadelphia, PA 19104-6021, *jshults@mail.med.upenn.edu*; Xiaoying Wu, University of Pennsylvania; Shiriki Kumanyika, University of Pennsylvania

Key Words: familial correlation structure, longitudinal study, quasi-least squares, health disparities, generalized estimating equations, correlated data

Longitudinal studies can address health disparities in diet related health conditions affecting minority populations. We consider a study that assesses

involvement of family or friends as a strategy for enhancing social support and success in weight loss in African Americans. The need to analyze family/friend clusters provides motivation for application of a familial correlation structure that previously was not available for generalized estimating equations (GEE). In general, by expanding the list of structures for GEE, we aim to improve efficiency, which could enable improved detection of clinically important differences by ethnic subgroup, or effects when the drop out rate is higher than expected, or the sample size is limited initially. This research was completed as part of an NIH funded program to address characteristics typical of longitudinal studies in minority populations.

Empirical Likelihood Method for Determining Nonparametric Spirometry Reference Values for Hispanic Americans

✱Nancy L. Glenn, Texas Southern University, Department of Mathematical Sciences, 3100 Cleburne St., Houston, TX 77004, *glennnl@tsu.edu*

Key Words: empirical likelihood, nonparametric, spirometry

The American Lung Association advocates regular pulmonary testing as it can detect lung diseases (such as cancer, chronic bronchitis, emphysema, and asthma) in their early stages and allow for timely treatment. Typical reference value equations employ linear regression forms that rely heavily on a parametric assumption that is often violated - normality. This violation may ultimately lead to misdiagnoses. We instead use empirical likelihood confidence intervals, a conceptually new nonparametric approach which replaces former reference value equations. Even though the United States has experienced substantial growth in the Hispanic population in the past few decades, there have not been any reference value studies specifically for Hispanic Americans. We use the empirical likelihood approach to obtain reference values for Hispanic Americans. This approach extends to other ethnicities.

375 Recent Development in Event History Data Analysis ●▲

SSC, Biopharmaceutical Section, WNAR
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Variance Estimation in Semiparametric Models with Censored Data

✱Zhezhen Jin, Columbia University, 722 W 168th St, New York, NY 10032, *zj7@columbia.edu*

Key Words: Estimating functiom, Monte Carlo Method, Censored data, Semiparametric model

It is challenging to obtain variance estimates if estimating functions are nonregular, such as nonsmooth and nonmonotone, in particular for censored data. We develop Monte Carlo smoothing method for variance estimation which might not require any explicit form, not even the estimating function itself. We present general convergence theory and demonstrate the methods with examples.

Modeling and Analyzing Data on Recurrent and Terminal Events

✱John Kalbfleisch, The University of Michigan, Department of Biostatistics, Ann Arbor, MI 48103, *jdkalbfl@umich.edu*

Key Words: recurrent events, estimating equations, marginal models

In clinical and observational studies, recurrent event data (e.g. repeated hospitalizations) are often encountered. In some important applications, the recurrent events are censored by a terminal event (e.g. death) that is strongly correlated with the recurrent events. We consider marginal models and methods that avoid the introduction of strong distributional assumptions about the recurrent event process. Methods based on estimating functions are used to assess dependence on covariates and to estimate the correlations between the recurrent and terminal processes. The approach is compared with others suggested in the literature and illustrated on data on recurrent hospitalizations and failure of treatment that arise in a Canada/USA study of peritoneal dialysis as a treatment for end stage renal disease.

Evaluation of Treatment Effects with Recurrent Events Under Dependent Censoring

✱Richard Cook, University of Waterloo, Waterloo, ON N2L 3G1 Canada, *rjcook@uwaterloo.ca*; Jerald Lawless, University of Waterloo

Key Words: recurrent events, tests of treatment effect, dependent censoring, gap times

In some clinical trials recurrent events are observed subject to dependent censoring. We consider trials involving two types of recurrent events with censoring rules based on one of these types. Under a mixed Poisson framework, a copula function is used to construct a bivariate gamma random effects distribution, which yields negative binomial marginal event processes. The performance of the Nelson-Aalen estimator and estimators of multiplicative treatment effects are examined under dependent censoring schemes, as well as estimates based on the proposed model. The methods are illustrated using data from a recent study in respirology.

376 Bayesian Methods for Medical and Epidemiologic Studies ●

WNAR, Section on Bayesian Statistical Science, Section on Statistics in Epidemiology, Biometrics Section
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

A Bayesian Hidden Markov Model for Motif Discovery Through Joint Modeling of Genomic Sequence and ChIP-Chip Data

✱Joseph G. Ibrahim, The University of North Carolina at Chapel Hill, Department of Biostatistics, McGavran-Greenberg Hall, Chapel Hill, NC 27599, *ibrahim@bios.unc.edu*; Jon A.L. Gelfond, University of Texas Health Science Center at San Antonio; Mayetri Gupta, Boston University

Key Words: DNA, Gene Regulation, Tiling Array, Sequence Motif

We propose a unified framework for the analysis of Chromatin (Ch) Immunoprecipitation (IP) microarray (ChIP-chip) data for detecting transcription factor binding sites (TFBSs) or motifs. Present analytical methods use a two-step approach: (i) analyze array data to estimate IP enrichment peaks then (ii) analyze the corresponding sequences independently of intensity information. The proposed model integrates peak finding and motif discovery through a unified Bayesian hidden Markov model (HMM) framework that accommodates the inherent uncertainty in both measurements. A Markov Chain Monte Carlo algorithm is developed, adapting recursive techniques used for HMMs. In simulations and applications to a yeast RAP1 dataset, the proposed method has favorable TFBS discovery performance compared to currently available two-stage procedures in terms of sensitivity and specificity.

Empirical Bayes-Type Shrinkage Estimation in Genetic Epidemiology

✱Bhramar Mukherjee, The University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48103, *bhramar@umich.edu*; Nilanjan Chatterjee, National Cancer Institute

Key Words: case-control studies, family-based studies, gene-environment interaction, power, type 1 error, Bayes

In analyzing data from case-control genetic association studies, plausible constraints on the exposure distribution like gene-environment independence, Hardy-Weinberg equilibrium are being exploited to yield efficient estimation strategies. However, these modern retrospective methods may incur bias under departures from the assumed constraints. The Bayesian paradigm is a natural alternative to build uncertainty around the assumed constraints, but comes with computational challenges, especially for large-scale association studies. We propose an empirical Bayes-type shrinkage estimation strategy for such problems, striking a balance between efficiency and robustness, boosted with simple computation. The problem of estimating gene-environment interaction from case-control data and that of combining related and unrelated controls in family based studies will serve as illustrative examples.

Applying Bayesian Ideas in a Multisite fMRI Study

✱Hal Stern, University of California, Irvine, *sternh@uci.edu*

Key Words: variance components, reliability, imaging, hierarchical models

Use of multiple sites in a brain imaging study provides an opportunity to increase sample size beyond what can typically be achieved at a single site. Multiple site studies however introduce a new source of variation, scanner-to-scanner variation, that can complicate analysis. The FBIRN (Function Biomedical Informatics Research Network) project has been exploring methods for multi-scanner studies in the context of schizophrenia. Bayesian ideas influence design of studies, reliability analysis, and data analysis.

377 Dynamic Factor Models for Real-Time Macroeconomic and Financial Forecasting ●

Business and Economics Statistics Section

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Modeling High-Dimensional Time Series by Generalized Linear Dynamic Factor Models

✱Manfred Deistler, Vienne University of Technology, Argentinierstr 8, Vienna, A 3100 Austria, *deistler@tuwien.ac.at*

Key Words: state space, factorization, rational spectra

We consider generalized linear dynamic factor models (GDFM) in a stationary context, where the observations are the sum of two uncorrelated processes: The so called latent process (obtained from a linear transformation of a low dimensional factor process) which has a singular rational spectral density and shows strong dependence and the noise process which has weak dependence in the cross-section thus generalizes the classical assumption of cross-sectionally uncorrelated noise. We present a structure theory consisting of factorization of rational singular spectra, a result on genericity of zeroless "tall" transfer functions, a special factorization and realization by state space systems, and a result on "averaging out" of noise for the cross-sectional dimension tending to infinity. Based on this a structure theory an estimation procedure is proposed.

Factor Decomposition of VARMA Models Based on Weighted Forecast-Error Covariances: Applied to Forecasting Quarterly GDP at Monthly Intervals

✱Peter Zadrozny, Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 3105, Washington, DC 20212, *zadrozny.peter@bls.gov*; Baoline Chen, Bureau of Economic Analysis

Key Words: principal components, Kalman filter, maximum likelihood, mixed-frequency data

We develop and apply weighted forecast-error covariance decomposition (WFECD) for reducing a large estimated VARMA data model of important variables of primary interest and unimportant auxiliary variables to a small VARMA factor model. WFECD has some advantages over principal components decomposition for developing parsimonious time-series models such as being able to handle nonstationary and rational-spectral models. We apply WFECD by estimating four monthly VARMA models of U.S. monthly coincident and leading indicators and quarterly real GDP. We WFECD decompose the estimated models, test for significant factors, compute small monthly significant-factor models, and compute and evaluate out-of-sample monthly GDP forecasts. The small monthly factor models produce monthly out-of-sample forecasts of quarterly GDP which are almost as accurate as those produced by the large data models.

Monthly Real-Time Estimates of Final GDP Based on an Estimated Monthly Model of Initial and Revised GDP

✱Baoline Chen, Bureau of Economic Analysis, 1441 L Street, NW, Washington, DC 20230, *baoline.chen@bea.gov*; Peter Zadrozny, Bureau of Labor Statistics

Key Words: Maximum likelihood model estimation, Real-time Kalman filtering, Data revisions and delays, Mixed-frequency data

The paper develops a two-step method based on an estimated VAR model for computing monthly estimates of quarterly GDP. Step 1 uses 3 initial GDP estimates released quarterly, 3 revised GDP released annually, and related variables released monthly, in historical form (indexed by periods to which they pertain), to compute ML estimates of a monthly VAR model. The 3rd annual GDP release is considered "final." Step 2 computes current monthly (filtered) estimates of the "final" GDP by applying the Kalman filter to the estimated model and the data in real-time form (indexed by periods in which they are observed). Because the estimated model is monthly, the filtered estimates of "final" GDP are produced at monthly intervals. The estimates are evaluated for RMSE accuracy relative to the 3rd-annual "final" GDP releases.

Nowcasting UK GDP Growth: An Evaluation of Dynamic Factor Models Using Quasi Real-Time Data

✱Jana Eklund, Bank of England, Threadneedle Street, London, EC2R 8AH United Kingdom, *Jana.Eklund@bankofengland.co.uk*; Kamath Kishore, Bank of England; Vincent Labhard, European Central Bank; Simon Price, Bank of England

Key Words: Kalman filtering, maximum likelihood, latent variables, mixed-frequency data

We evaluate a range of models for nowcasting UK GDP growth. Nowcasting (estimating current data before they are published using information currently available) is crucial for macroeconomic policymakers. Many methods exist to perform this task, but dynamic factor models are particularly well suited for it, because they efficiently bring a wide range of data to this problem, including mixed-frequency data. An important issue in evaluating such models is to take account of data revisions and the precise timing of releases, which has a major impact on the information available at any point. Practical difficulties prevent us from constructing a complete real-time data set, but historical real-time data are available for GDP growth and other key variables, and information about timing of data releases is incorporated in the "quasi-real-time" data set we use.

NOWcasting Quarterly German GDP at Monthly Intervals Using Dynamic Factor Models

✱Klaus Wohlrabe, Ifo Institute for Economic Research, Poschingerstrasse 5, Munich, 81679 Germany, *wohlrabe@ifo.de*

Key Words: principal components decomposition, Kalman filtering, latent variables, mixed-frequency data

The paper applies Stock and Watson's combination of principle components decomposition and the EM algorithm (*JBES*, 2002) to German quarterly GDP and other German monthly macroeconomic variables in order to compute monthly factors and monthly estimates of quarterly GDP. Then, the paper uses the computed monthly factors and estimated GDP to estimate monthly models for computing and evaluating out-of-sample monthly forecasts of quarterly German GDP.

378 Predicting Cancer Risk Using Flexible Methods ●▲

Section on Risk Analysis, Section on Statistics in Epidemiology, WNAR, Section on Health Policy Statistics

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Risk Prediction in Hereditary Cancer Syndromes

✱Sining Chen, Johns Hopkins University, 615 N. Wolfe St., W7041, Baltimore, MD 21205, *sining@jhmi.edu*; Giovanni Parmigiani, Johns Hopkins University; Yue Yin, Genentech, Inc.; Wenyi Wang, University of California, Berkeley; Patrice Watson, Creighton University; Henry T. Lynch, Creighton University

Key Words: genetic risk

We discuss the estimation of genetic risk in hereditary cancer syndromes. Because inherited mutation in major genes are rare in the population, the commonly available data sets are mainly multiple-case families. Using large families with hereditary colon cancer as example, We compare a Kaplan-Meier analysis on the prospective component of the data set, and a conditional likelihood-based method using both prospective and retrospective components.

A Colorectal Cancer Risk Prediction Tool

✱Ruth Pfeiffer, National Cancer Institute, Biostatistics Branch, Div. of Cancer Epidemiology and Genetics, 6120 Executive Blvd, Bethesda, MD 20892-7244, *pfeiffer@mail.nih.gov*

Key Words: Absolute risk, Family history, Screening, Counseling, Risk Prediction

Several modifiable risk factors have been identified for colorectal cancer (CRC), the second leading cause of cancer death in the US. We developed models to estimate the probability of first incidence of proximal, distal or rectal cancer in White men and women, aged 50+ years. We 1) estimated relative risk parameters from population-based case-control data separately for proximal, distal, and rectal cancer; 2) estimated baseline age-specific cancer hazard rates based on SEER incidence rates and attributable risks; and 3) combined competing risks from national mortality rates, relative risks and baseline hazards to estimate the probability of developing the first of proximal, distal or rectal cancer over a pre-specified time interval (e.g., 5 or 10 years) given age and risk factors. We derived the variance of the risk estimates. Relative risks differed by gender and tumor site.

Building Cancer Risk Models Using Multifarious Complex Data Sources: Exploiting the Elegance of Bayes

✱Donald A. Berry, M.D. Anderson Cancer Center, Dept of Biostatistics, 1515 Holcombe Blvd., Houston, TX 77030, *dberry@mdanderson.org*

Key Words: Bayesian modeling, Bayesian calculations, Complex population, Breast cancer screening effectiveness, Breast cancer treatment effectiveness

Bayesian calculations are neat and tidy when one can write down a mathematical function expressing the likelihood. But what to do when in complicated settings where there is no such likelihood? An example is addressing the effects of a various interventions in a population. I will present methodology used as a part of a highly publicized report concerning reductions in breast cancer mortality observed in the US in the 1990s. The data used included incidence by stage of disease as reported by SEER and mortality as per the NCHS. Other databases were used for patterns of screening and treatment and their efficacies. I will discuss the interplay between assessing prior distributions (for 6 parameters) and incorporating information when there is no explicit likelihood. Simulations and the rejection method yield posterior distributions of the benefits of screening and treatment.

379 A New Paradigm of Statistical Data Analysis: 'Omics' Data ●▲

Biometrics Section, Biopharmaceutical Section, WNAR

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Reverse Engineering To Construct Protein-Protein Interaction Network

✱Susmita Datta, University of Louisville, Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, 555 South Floyd St., Louisville, KY 40202, *susmita.datta@louisville.edu*; Vasyl Pihur, University of Louisville; Somnath Datta, University of Louisville

Key Words: Gene, Protein, Network, PLS

Detecting gene-gene and protein-protein interactions is an expensive and labor-intensive process. The two hybrid systems are the most common method used for detecting physical interactions. Unfortunately, due to its high false positive discovery rates, it has received a lot of criticism in the literature. Hence computational approaches to reverse engineering of gene and protein association networks are of great interest to biologists as they allow for an indirect elucidation of relationships between genes or proteins through existing post-gnomic, usually microarray data. Examples of computational approaches include Bayesian networks, auto-regressive models, correlation-based models, clustering techniques, etc. We propose a novel computational method based on the partial least squares to uncover the relationships between genes/proteins.

Using ‘Intelligent Systems’ for Predicting Protein Function and Selecting Targets for Structural Genomics

✱Rajesh Nair, Columbia University, 1130 St. Nicholas Ave, 8th. floor, Room 805, New York, NY 10032, *rajnair5@gmail.com*; Burkhard Rost, Columbia University

Key Words: bioinformatics, protein structure, protein function, target selection, structural genomics, function prediction

Large-scale sequencing projects have resulted in a flood of protein sequence data. Two major challenges in reducing the rapidly widening gap in our knowledge of the structure and function of proteins are i) an effective sampling strategy for identifying candidates for structure and function determination and ii) tools for accurately inferring protein structure and function from what is known. I will discuss some of the strategies we have pursued to tackle these challenges. In the context of identifying candidates for structure determination, I will discuss strategies for target selection and prioritization that we have implemented to reduce the cost per structure. Application of these strategies has resulted in over 300 protein structures solved at the North Eastern Structural Genomics (NESG) consortium, most of which provided novel structural information for large-sequence families.

Regularization and Variable Selection for Data with Interdependent Structures

✱Jun Xie, Purdue University, 250 N. University Street, West Lafayette, IN 47907, *junxie@stat.purdue.edu*; Lingmin Zeng, Purdue University

Key Words: Elastic net algorithm, Grouping effect, LARS algorithm, Ridge regression, Variable selection

Variable selection methods are powerful tools for analyzing high dimensional massive data. In bioinformatics, the methods have been applied in gene expression microarray data analysis. It is well-known that for genes sharing a common biological pathway or a similar function, the correlations among them can be very high. However, most of the available variable selection methods cannot deal with complicated interdependence among data. We propose two new algorithms, namely gLars and gRidge, to select groups of highly correlated variables together in regression models. The new approaches intent to conduct grouping and selecting at the same time. Simulations and a real example show that our proposed methods often outperform the existing variable selection methods, including LARS and elastic net, in terms of both prediction error and preserving sparsity of representation.

Statistical Proteomics Initiative: Research Initiative for Proteomics Sequence Analysis

✱Martin McIntosh, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave East, M2-B820, seattle, WA 98109, *mmcintosh@fhcrc.org*; Olga Vitek, Purdue University

Key Words: Tande MS, Signal Processing, Sequence Analysysis

Like other biotechnologies have done for genomics and transcriptomics, revolutions in mass spectrometry instrumentation, computation, and modern biological databases are driving high throughput biological inquiry in protein sequence discovery and analysis. The need and opportunities for statistical research using quantitative tandem mass spectrometry (Tandem MS, or MS/MS) are far greater than that of perhaps any other -omics data types. To foster advancements the United States Human Proteome Organization (USHUPO) is developing a Statistical Proteomics Initiative to advance quantitative research in this area. In this talk we outline the initiative, and also several core research questions facing the field, including presenting several data sets which researchers may wish to use to help begin research in this field.

380 *The Black Swan*: A Discussion ●▲

The American Statistician
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

The Black Swan: A Discussion

✱Aaron Brown, AQR Capital Management, *Aaron.Brown@privateeram.com*; ✱Robert Lund, Clemson University, *lund@ces.clemson.edu*; ✱Nassim N. Taleb, Self Employed, *nnt@fooledbyrandomness.com*; ✱S. Stanley Young, National Institute of Statistical Sciences, RTP, NC 27709, *young@niss.org*; ✱Donald B. Rubin, Harvard University, Department of Statistics, 1 Oxford Street, 7th Floor, Cambridge, MA 02138, *rubin@stat.harvard.edu*

Key Words: financial analysis, outliers, philosophy, probability model, rare events, statistical model

The book *The Black Swan: The Impact of the Highly Improbable*, by Nassim Nicholas Taleb, was released with much acclaim in 2007. The book was highly critical of statisticians and quantitative social scientists (“quants”) with respect to their abilities to model randomness, particularly with regard to the ability to predict rare, influential events—“black swans.” Several reviews appeared in the August 2007 edition of *The American Statistician*, with rebuttal by the author. This panel session will provide discussions of the book, of its impact and relevance, and of the applicability of probability models to rare events.

381 Paradata, Data Quality, and the National Health Interview Survey (NHIS) ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Analyzing Field Notes Systematically To Better Understand Respondent Participation

Barbara Stussman, National Center for Health Statistics; ✱Catherine Simile, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *csimile@cdc.gov*

Key Words: paradata, interactions, textual data, participation, field procedures

Paradata provide a unique opportunity to learn about interactions between interviewers and respondents in the National Health Interview Survey, conducted annually by the National Center for Health Statistics, Centers for Disease Control and Prevention. Interviewers record textual entries on a variety of topics including reasons for partial and break-off interviews, reasons for telephone interviews (given that the NHIS is primarily a face-to-face survey), and general information about the family’s availability, cooperativeness, and circumstances surrounding participation. Using qualitative and quantitative analysis, this paper will examine textual entries recorded in the 2004–2007 NHIS. The paper will describe the process of analyzing textual data, present findings, and explain how the findings have led to changes in the collection of paradata and modifications to field procedures.

Do You Really Mean What You Say? Doorstep Concerns and Data Quality in the National Health Interview Survey (NHIS)

✱James M. Dahlhamer, National Center for Health Statistics, 3311 Toledo Road, Room 2117, Hyattsville, MD 20782, *jdahlhamer@cdc.gov*; Catherine Simile, National Center for Health Statistics; Beth Taylor, National Center for Health Statistics

Key Words: paradata, data quality, satisficing, measurement error

Measurement error arising from respondent-interviewer interactions during the interview has received considerable attention among survey researchers. Less understood are the effects on data quality of respondent concerns and reluctance expressed during initial doorstep interactions. While some suggest that doorstep concerns foretell a respondent's level of commitment to the interview, others suggest that stated concerns carry little intrinsic meaning. Using paradata (data about the collection process) and health data from the 2006 and 2007 NHIS, we explore the associations between respondent doorstep concerns and reluctance (e.g., "too busy") and various measures of data quality. We also attempt to extend prior research by exploring the effects of combinations of respondent concerns. We discuss the implications of our findings for interviewing and quality assurance procedures.

The 2006 National Health Interview Survey Paradata File: Overview and Applications

✱Beth Taylor, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *bft8@cdc.gov*

Key Words: paradata, interviewer strategies, partial, cooperation, contactability, mode

In January 2008, the National Health Interview Survey released its first annual Paradata File, containing information about the data collection process for the 2006 NHIS. Analyses of paradata can be used to explore case characteristics and allow for a better understanding of NHIS respondents and the interviewer-respondent dynamic. This discussion will provide an overview of the 2006 NHIS Paradata file, including the scope of the cases included on the file and a summary of major variables related to interviewer strategies, measures of contactability and cooperation, measures of time, mode of interview, and reasons for partial interviews and break-offs. Applications of the data as a stand-alone file and as a file linked to the 2006 NHIS health data files will also be discussed, as well as examples of past research using paradata from the NHIS.

An Exploration into the Use of Paradata for Nonresponse Adjustment in a Health Survey

✱Aaron Maitland, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD , *amaitland@cdc.gov*; Carolina Casas-Cordero, University of Maryland; Frauke Kreuter, University of Maryland

Key Words: Nonresponse, Paradata

Nonresponse is a growing concern to survey researchers. One strategy for reducing the effect of nonresponse is statistical adjustment of survey estimates. Methods available for nonresponse adjustment include post-stratification and ratio adjustment. Another approach starts with populating the sample frame with a set of auxiliary variables that are correlated with both the probability of response and the survey variables of interest. The recently released 2006 National Health Interview Survey paradata file contains information, collected during the process of recruiting sample households to participate, that could be potentially useful for nonresponse adjustment. Information collected includes indicators of respondent reluctance and strategies for recruiting sample members. The goal of this paper is to test hypotheses about correlations between paradata variables and survey outcomes.

382 What Can SPAIG Do for You? ▲

SPAIG Committee, Section on Physical and Engineering Sciences
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

SPAIG Activities at Iowa State University

✱Dean L. Isaacson, Iowa State University, Department of Statistics, Iowa State University, Ames, IA 50011, *dli@iastate.edu*

The Department of Statistics at Iowa State University started having formal partnerships with industry in 1993 when General Motors asked that the Department offer the MS degree to their employees. This initial partnership led to many other formal and informal partnerships with Business, Industry and Government. This talk will describe the benefits and costs that resulted from these interactions over the past 15 years. The benefits were many so when the Department submitted a proposal for a Research Training Grant in 2005, industry and government partnerships were a central part of the proposal. The majority of this talk will focus on the goals of the RTG grant and the successes obtained thus far. The goals of SPAIG will be apparent in what was proposed and what has been done through this RTG grant.

University-Industry Demonstration Partnership

Robert Starbuck, Wyeth Research; ✱Anthony Boccanfuso, The National Academy of Sciences, University-Industry Demonstration Partnership, 500 5th St NW, Keck Building 550, Washington, DC 20001, *ABoccanfuso@nas.edu*

Key Words: partnerships, research, UIDP, competitiveness, innovation, intellectual property

For many universities, industry-sponsored research is a critical funding source. Federal legislation, such as the American Competitiveness Initiative, also recognize its value. Recent surveys, however, show that growth in industry-sponsored research at US universities has stagnated and research specifically linked to licenses and options has actually decreased. Thus the University-Industry Demonstration Partnership (UIDP), whose mission is to nourish and expand collaborative partnerships between university and industry in the US, was formed. The UIDP is engaged in collaborative and disruptive experiments on new approaches to sponsored research, licensing arrangements, and the broader strategic elements of a healthy, long-term university-industry relationship. This presentation will describe the work of the UIDP, results from recent efforts, and its planned initiatives.

Summary of Recent Salary Surveys of Statisticians

✱Boris Iglewicz, Temple University, Department of Statistics, 1810 North 13th Street, Philadelphia, PA 19122, *borisi@temple.edu*

Key Words: academic, government, industry, inflation adjusted, non-response, salary surveys

The American Statistical Association and related professional societies have sponsored periodic salary surveys of statisticians. The ASA web site contains summaries of such salary surveys, including the 2007 SPAIG survey of industry and government statisticians, plus annual surveys of academic statisticians and also of biostatisticians. In addition, the American Mathematical Society has been running annual salary surveys of mathematicians and statisticians. A number of other professional organizations also run salary surveys of their members. A summary of data from these surveys will be provided with results compared. Also considered will be inflation adjusted salaries over time and comparisons with comparable salaries of professionals from other fields.

Strategic Partnerships and the ASA

✱Ron Wasserstein, The American Statistical Association, *ron@amstat.org*

The American Statistical Association has a rich heritage of membership across the three broad sectors of academe, industry, and government. This breadth provides strength to the association, but presents challenges, as well. Issues facing the association, and directions for the future, will be examined in this context by the ASA's executive director.

Partnerships with Nonprofit Research Institutions

✱Sally C. Morton, RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709, *morton@rti.org*

Key Words: Alliance, Partnership, SPAIG

Partnering with a nonprofit research institution can be mutually beneficial. How can your organization form an alliance with a nonprofit institution to create a "win-win" situation for both entities? The missions and objectives of nonprofits will be discussed and how partnering can contribute to those objectives. Topics will include how financial and other incentives impact decisions about partnering. The discussion will also encompass how organizational and management structures might influence the architecture of a partnership. Well-designed and realistic partnerships are a means by which to expand your organization's capabilities, potential, and value.

383 Novel Analyses of Large Nonrandomized Data Sets To Improve Policymaking in Health and Medicine ●

Section on Health Policy Statistics, Section on Statistical Computing, Social Statistics Section

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Geographic Area Variations in the Medicare Health Plan Era

✱Patricia Keenan, Yale School of Public Health, 60 College St, PO Box 208034, New Haven, CT 06520, *patricia.keenan@yale.edu*; A. James O'Malley, Harvard Medical School; Bruce Landon, Harvard Medical School; Paul Cleary, Yale School of Public Health; Lawrence Zaborski, Harvard Medical School; Alan Zaslavsky, Harvard University Medical School

Key Words: quality

We examined variation in Medicare beneficiary care assessments across geography, fee-for-service (FFS) or Medicare Advantage (MA) private plan type, and individual MA plans using CAHPS surveys. A hierarchical model with state, health referral region (HRR), plan, and person levels estimated correlated variance components for FFS and MA random effects at state and HRR levels and for MA plan within HRR, linearly adjusting for case-mix. We found greater geographic variation in MA than in FFS for 10 of 11 measures. Geographic effects on FFS and MA scores were strongly correlated at both geographic levels. In all measures, MA plan variation within HRRs accounts for > 1/4 of the combined variance across HRRs/states. Medicare private plan enrollment, now 1/5 of beneficiaries, may widen variations in care experiences even as other policy efforts aim to reduce existing geographic variations.

Selection into Medicare Coverage: Propensity Scores, Multiple Treatments, and Comparisons of Beneficiary Evaluations of Care

✱Amelia Haviland, RAND Corporation, 4570 Fifth Avenue, Pittsburgh, PA 15213, *haviland@rand.org*; Marc N. Elliott, RAND Corporation; Katrin Hambarsoomians, RAND Corporation

Key Words: Multinomial, Prescription Drugs, Consumer Evaluation of Healthcare, Observational Data, Propensity Score

Medicare beneficiaries may choose among many coverage options that target different beneficiaries and differ in costs to both beneficiaries and taxpayers; understanding any differences in healthcare experiences by plan type may be important in establishing which best serve seniors. In this paper, we use the 2007 Medicare CAHPS survey to compare beneficiary experiences by plan type. Because beneficiaries are not randomly assigned, we compare estimates of beneficiary experience using regression to control for differences in enrollee characteristics to estimates from a multiple-groups propensity score analysis. We estimate average treatment effects and effects of treatment on the treated and demonstrate that the "optimal" coverage varies by enrollee characteristics. Finally, we consider bias/variance tradeoffs in propensity techniques and the role of sample size.

Comparing Additive and Hierarchical Classifications of Race/Ethnicity with Respect to Evaluations of Care

✱David Klein, RAND Corporation, 1776 Main St., Santa Monica, CA 90401, *dklein@rand.org*; Marc N. Elliott, RAND Corporation; Amelia Haviland, RAND Corporation

Key Words: Multiracial, Health Disparity, Native American

Self-reported race/ethnicity is vital in measuring health disparities. The standard items for eliciting race/ethnicity ask first about Hispanic ethnicity and then allow respondents to check one or more options from a list of 6 racial groups. Standard approaches apply hierarchical classifications (HC), create heterogeneous "multiracial" groups, or attempt to impute the single choice that would have been made if only one were allowed. We use a large Medicare CAHPS dataset to compare simple HC to (1) an additive model of race/ethnicity, and (2) an all-possible-combinations approach in capturing racial/ethnic variation in healthcare evaluations. Additive models produce similar estimates for large groups (White, Black, Asian, Hispanic), but may more accurately capture the experiences of Native American and multiracial respondents than HC.

Improving Subgroup Comparisons of Consumer Reports by Adjusting for Differences in Extreme Response Tendency

✱Marc N. Elliott, RAND Corporation, 1776 Main St, Santa Monica, CA 90401, *elliott@rand.org*; Amelia Haviland, RAND Corporation; David Kanouse, RAND Corporation; Katrin Hambarsoomians, RAND Corporation

Consumer health care evaluations are vital to understanding racial/ethnic and other health disparities. Previous patient surveys have found counter-intuitive patterns, even after standard case-mix adjustment (CMA)- e.g. less positive ratings for those with higher income, more positive evaluations by Blacks and Hispanics than Whites. Using 0–10 Medicare CAHPS ratings, we show that in addition to the positive response tendency modeled by CMA, there is also an extreme response tendency (ERT) that varies among beneficiaries, with some subgroups less likely to use both extremes of the scale. With negatively skewed healthcare evaluations, ERT may be misinterpreted as more positive healthcare experiences in the absence of specific adjustment. We derive such an adjustment, with substantial effects on disparity estimates, resolving or attenuating many counterintuitive findings.

Comparing Traditional Instrumental Variables Estimation to Parametric Modeling of a Problem in Health Policy

✱A. James O'Malley, Harvard Medical School, Dept of Health Care Policy, 180 Longwood Avenue, Boston, MA 02115, *omalley@hcp.med.harvard.edu*; Sharon-Lise Normand, Harvard Medical School; Richard Frank, Harvard Medical School

Key Words: Antipsychotic drug, Instrumental variable, Multivariate model, Orthogonality, Schizophrenia

We provide a statistical framework based on instrumental variables methods for assessing whether new antipsychotic drugs for treatment for schizophrenia have a lower net cost than conventional drugs. The data are analyzed using standard non-parametric instrumental variables estimators that do not make parametric assumptions and also using a parametric model that relies on bivariate normality. We find surprisingly large differences in the results between the methods. To investigate the differences, we performed a series of simulation experiments in which the control condition generates data that looks like the observed data. The results of these experiments yielded several interesting properties of the methods and allowed rules of thumb for practical implementation to be established.

384 Statistical Issues of Thorough QT/QTc Studies in Clinical Trials ●▲

Biopharmaceutical Section, Biometrics Section
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

New Methods on QT Correction

✱Bin Cheng, Columbia University, *bc2159@columbia.edu*

Key Words: QT

The current procedure for QT correction typically assumes a linear or loglinear relationship between QT and heart rate measures and heart rate is treated as deterministic rather than as random itself. In this paper, we propose a flexible correction method based on nonparametric smoothing and measurement error model. Simulation is conducted to compare the proposed method with the existing procedures. Data from several thorough QT/QTc trials are reanalyzed.

Exploration of Covariance Structure for TQT Study

✱Xiaodong Li, Merck & Co., Inc, *li_xiaodong@merck.com*

Key Words: TQT, covariance, Toeplitz, compound symmetry, AIC, BIC

QT interval represents the time between the start of ventricular depolarization and the end of ventricular repolarization (QT). QTc is the QT-interval corrected for heart rates. QT prolongation is dangerous and may cause death in some cases.Current FDA ICH E 14 guidance requires all sponsors submitting new drug applications to conduct a thorough QT/QTc study to demonstrate the lack of QT prolongation effect for their drug.The primary endpoint is the time-matched mean difference between the drug and placebo after baseline adjustment at each time point. Paired t-tests are often used by regulatory agencies but a linear model may be more appropriate. The appropriate covariance structure for a linear model approach is still in debate. This talk will focus on the results of an analysis which shows similar patterns in the covariance matrices for different compounds.

Statistical Issues of QT Prolongation Assessment Based on Linear Concentration Modeling

✱Yi Tsong, U.S. Food and Drug Administration, CDER, 5244 10903 New Hampshire Ave, Silver Spring, MD 20993, *yi.tsong@fda.hhs.gov*; Meiyu Shen, U.S. Food and Drug Administration; Zhong Jinglin, U.S. Food and Drug Administration; Joanne Zhang, U.S. Food and Drug Administration

Key Words: ICH, Thorough QTc, Concentration-response, Linear model, Intersection-union test

The ICH E14, 2005 defined that drug-induced prolongation of QT interval as evidenced by an upper bound of the 95% confidence interval around the mean effect on QTc of 10 ms or higher. This objective leads to the application of intersection-union test by testing the mean difference between the treatment and placebo of QTc change from baseline at each of the matched time point at which the observations are collected. The multiple comparison requirement nature of the intersection-union test leads to the concern of study efficiency. Based on the concept of clinical pharmacology, a concentration-response modeling approach was often adopted to assess the prolongation size of QTc interval induced by drug without carefully examine the validity of the assumptions involved. In most of the applications, the model is assumed to be either linear, log-linear or logistic.

385 Causal Inference: Advances and Applications

Biometrics Section, Section on Nonparametric Statistics, Biopharmaceutical Section, Section on Bayesian Statistical Science, Section on Health Policy Statistics
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Semiparametric Estimation and Inference for Distributional and General Treatment Causal Effects

✱Jing Cheng, University of Florida, 1329 SW 16th Street, Room 5130, Division of Biostatistics, Gainesville, FL 32610, *jcheng@biostat.ufl.edu*; Jing Qin, National Institute of Allergies and Infectious Diseases; Biao Zhang, The University of Toledo

Key Words: Causal effect, empirical likelihood, exponential tilt model, instrumental variable, noncompliance, randomized trials

There is a large literature on methods of analysis for randomized trials with noncompliance by focusing on the causal effect of treatment on the average outcome. In this paper, we develop an instrumental variable approach by using empirical likelihood method with the exponential tilt model to estimate and conduct inference on the treatment effect on the entire outcome distributions and their general functions for compliers under treatment and control. Our method can be applied to general outcomes. The method is illustrated by an analysis of data from a randomized trial of an encouragement intervention to improve adherence to prescribed depression treatments among depressed elderly patients in primary care practices.

Understanding Mediation Using Principal Stratification

✱Michael R. Elliott, The University of Michigan, Department of Biostatistics, M4041 SPH II, School of Public Health, 1420 Washington Heights, Ann Arbor, MI 48109, *mrelliot@umich.edu*; Trivellore Raghunathan, The University of Michigan

Key Words: causal modeling, direct effects, sequential randomization

Mediation analysis investigates whether the effect of a treatment goes through or around a post-treatment mediator. Because mediators are observed after treatment assignment, standard analyses require the sequential randomization assumption. We develop a principal stratification approach (Frangakis and Rubin 2002) that relaxes sequential randomization by considering counterfactual distributions of mediator and outcome. We define direct effect principal strata by subjects whose mediator is unchanged under the treatment assignments, and consider the direct effect as the intent-to-treat relative risk within the direct effect principal strata. While our models are underidentified for dichotomous outcomes, inferential statements are still possible because of boundary conditions imposed by the data. Bayesian approaches allow incorporation of a weak monotonicity assumption to sharpen inference.

An Integrated Nonparametric Approach to Missing Data in Causal Inference

✱Jennifer Hill, Columbia University, 420 W. 118th St. Rm 1512, New York, NY 10027, *jh1030@columbia.edu*; Robert McCulloch, The University of Chicago

Key Words: causal inference, missing data, nonparametric, Bayesian, pattern mixture, BART

Use of data mining algorithms has been proposed in recent years as a flexible approach to the causal inference problem under the assumption of an ignorable treatment assignment mechanism. The advantages of one such algorithm, Bayesian Additive Regression Trees (BART), have been demonstrated in earlier work. This approach handles outcome missing data fluidly. New work, relying on a pattern mixture approach to missing data, proposes an extension to the model and MCMC algorithm to accommodate predictor missing data as well. Ease of implementation and desirable properties are demonstrated through simulations and an example.

Deconfounding Small Quasi-Experiments Using Propensity Scores and Other Dimension Reduction Techniques

✱Yevgeniya Kleyman, The University of Michigan, 605 Hidden Valley Club Dr., Apt 204, Ann Arbor, MI 48104, *ykleyman@umich.edu*; Ben Hansen, The University of Michigan

Key Words: observational study, propensity score, covariate balance, matching, prognostic score, full matching

When the sample is small and the number of potential confounders large, propensity adjustment may seem to have little to offer. However, in combination with dimension reduction, flexible matching, formal diagnostics and simple post-matching adjustments, it works surprisingly well. In our motivating example, a richly observed quasi-experiment comparing a faith-based and a conventional substance abuse treatment, close propensity score matching was not feasible. Still, by matching relatively coarsely on the propensity score but also matching on other scores summarizing the covariate, it was possible to balance k=27 covariates with only n=67 subjects; closer propensity matching was no better in these terms, and worse in others. In a thorough simulation study, we find that our combination of techniques reduces bias and Type I error rates in comparison to those resulting from rival approaches.

The Malaria Attributable Fraction: Definition, Inference, and Sensitivity Analysis

✱Dylan Small, University of Pennsylvania, 400 Huntsman Hall, 3730 Walnut St., Philadelphia, PA 19104, *dsmall@wharton.upenn.edu*

Key Words: causal inference, sensitivity analysis, malaria, attributable fraction

An important measure of the burden of malaria is the proportion of children that have fevers attributable to malaria, called the malaria attributable fraction (MAF). A difficulty in estimating the MAF is that it is hard to diagnose a fever as being due to malaria parasites as opposed to other illnesses. Microscopic examination of blood for parasites helps, but children living in areas of high malaria infelicity often tolerate parasites without the parasites causing malaria disease. We consider estimation of the MAF based on fever incidence and parasite density data. We present a potential outcomes framework for defining the MAF. The classical estimator depends on an assumption that parasite densities are effectively randomly assigned. We provide evidence that this assumption does not hold and develop a sensitivity analysis for the effect of departures from this assumption.

386 Recent Development of Bayesian Methods for Missing Data ●▲

Section on Bayesian Statistical Science, Section on Survey Research Methods, Section on Health Policy Statistics
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Diagnostic Measures for Generalized Linear Models with Missing Covariates

✱Hongtu Zhu, The University of North Carolina at Chapel Hill, 3108-B McGavran Greenberg Hall, CB#7420, Department of Biostatistics, Chapel Hill, NC 27599, *hzhu@bios.unc.edu*; Joseph G. Ibrahim, The University of North Carolina at Chapel Hill; Xiaoyan Shi, The University of North Carolina at Chapel Hill

Key Words: conditional residual, generalized linear models, case-deletion measure, missing data, residual process

In this paper, we carry out an in-depth investigation of diagnostic measures for assessing the influence of observations and model misspecification in the presence of missing covariate data for generalized linear models. Our diagnostic measures include case-deletion measures and conditional residuals. We use the conditional residuals to construct goodness of fit statistics for testing possible misspecifications in model assumptions, including the sampling distribution. We develop specific strategies for incorporating missing data into goodness of fit statistics in order to increase the power of detecting model misspecification.

An Exact Noniterative Sampling Procedure for Discrete Missing Data Problems

✱Guo-Liang Tian, University of Marylan Greenebaum Cancer Center, 10 South Pine Streect, MSTF Suite 261, Baltimore, MD 21201, *gtian2@umm.edu*; Kai Wang Ng, The University of Hong Kong; Ming T. Tan, University of Maryland Greenebaum Cancer Center

Key Words: Contingency tables, Data augmentation algorithm, EM algorithm, Gibbs sampler, IBF sampler, Markov chain Monte Carlo

Many statistical problems can be formulated as discrete missing data problems (MDPs). Examples include change-point problems, capture and recapture models, sample survey with non-response, zero-inflated Poisson models, diagnostic tests, bioassay and so on. This paper proposes an exact non-iterative sampling algorithm to obtain iid samples from posterior distribution in discrete MDPs. The algorithm is essentially a conditional sampling, thus completely avoiding problems of convergence and slow convergence in iterative algorithms such as Markov chain Monte Carlo. The key idea is to first utilize the sampling-wise inverse Bayes formula to derive the conditional distribution of the missing data given the observed data, and then to draw iid samples from the complete-data posterior distribution. We apply the method to contingency tables with one supplemental margin for an HIV study.

Bayesian Variable Selection for Modeling Repeated Binary Responses and Time-Dependent Missing Covariates

✱Fang Yu, University of Nebraska, Department of Biostatististics, 984350 Nebraska Medical Center, Omaha, NE 68198-4350, *fangyu@unmc.edu*; Ming-Hui Chen, University of Connecticut; Lan Huang, National Cancer Institute; Gregory J. Anderson, University of Connecticut

Key Words: Generalized Linear Mixed Model (GLMM), Missing at Random, MCMC Algorithm, Flower Intensity, Weather Conditions, Tilia

We develop a novel modeling strategy for analyzing data with repeated binary responses over time as well as time-dependent missing covariates. Covariates are assumed to be missing at random (MAR). We use the generalized linear mixed logistic regression model for the repeated binary responses and a joint model for time-dependent missing covariates using information from different sources. A new Bayesian method is developed to identify the importance of each covariates and the sensitivity of the specification of the missing covariates models is investigated. A Markov chain Monte Carlo algorithm is developed for computing the Bayesian estimates. A real plant dataset is used to motivate and illustrate the proposed methodology.

A Joint Modeling Approach to Repeated Measures of Mixed Data Types

✱W. John Boscardin, University of California, Los Angeles, 51-254 Center for Health Sciences, Box 177220, Departments of Biostatistics and Medicine, Los Angeles, CA 90095-1772, *jbosco@ucla.edu*; Xiao Zhang, The University of Alabama at Birmingham; Tom Belin, University of California, Los Angeles

Key Words: multiple imputation, multinomial probit, multivariate probit, restricted covariance matrix

From a Bayesian perspective, we propose a unified model for analyzing mixtures of continuous, ordinal and nominal repeated measures. The continuous measures are linked to latent variables for the ordinal and nominal measures via a multivariate normal model. MCMC inference is used with a step for sampling the restricted covariance structure. We will examine the application of this model in an incomplete data setting where our Bayesian implementation allows for simultaneous imputation and inference. Alternatively, separation of the imputer and analyst roles may be preferable in some situations. We compare our approach to various alternatives, including the general location model and switching regressions, in the context of a longitudinal study on functional outcomes of hospitalized older medical patients.

Bayesian Structural Equations Models for Longitudinal Surveys Data with Missing Responses and Covariates

✱Ming-Hui Chen, University of Connecticut, Department of Statistics, 215 Glenbrook Road, U-4120, Storrs, CT 06269, *mhchen@stat.uconn.edu*; Sonali Das, The Council for Scientific and Industrial Research; Sung Duk Kim, National Institute of Child Health and Human Development; Nicholas Warren, University of Connecticut Health Center

Key Words: Latent variables, Markov chain Monte Carlo, Missing at random, Random effects, VHA all employee survey data

Motivated by large survey data conducted by the U.S. Veterans Health Administration (VHA), we propose new methods for the analysis of longitudinal survey data using structural equations models (SEMs). A reparameterized version of the longitudinal SEM is developed to ensure model identifiability and facilitate efficient Bayesian posterior computation for large survey data. Several structural equations models are considered and compared via a variation of the deviance information criterion to investigate the importance of facility and covariate effects in the presence of missing data for the VHA study. We also propose an innovative model for the missing covariates for longitudinal survey data when no individual respondents can be traced over the survey period under investigation. A detailed analysis of three VHA all employee survey data is presented to illustrate the proposed methodology.

387 L-Moments: Recent Developments in Theory and Applications

Section on Nonparametric Statistics, IMS

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Parameter Estimation for General Univariate Distributions Using the Method of L-Moments

✱Jonathan Hosking, IBM T. J. Watson Research Center, PO Box 218, 1101 Kitchawan Road, Yorktown Heights, NY 10598, *hosking@watson.ibm.com*

Key Words: L-moments, distribution, parameter, estimate, R

Distribution parameters can be estimated by equating sample estimates of L-moments to their population analogues. When the distribution parameters can be expressed as functions of the population L-moments, sample L-moments can be plugged into these expressions to obtain parameter estimates. Explicit formulas or iterative algorithms are available for many commonly used distributions, but there are many other distributions for which no usable procedure has been obtained. A general solution is to compute L-moments given the parameters by numerical integration, and to compute parameters from L-moments by numerical optimization. Care is needed to avoid numerical difficulties, but we have developed a practical method that gives accurate results for a wide range of distributions. We describe the procedure and its implementation in the R language. Several illustrative examples are included.

Algorithms and Results for a CPU-Intensive, L-Moment-Based Approach Using R for Regionalized Flood Estimation in Texas

✱William H. Asquith, U.S. Geological Survey, 8027 Exchange Drive, Austin, TX 78754, *wasquith@usgs.gov*

Key Words: L-moments, Distributions, R environment, Floods, Wakeby, weighted-least squares

Flood estimates at ungaged sites are vital to infrastructure design. Estimates can be made by PRESS-minimized, weighted-least squares, multilinear regression (WLS) between the flood distribution at gaged sites (stations) and watershed features such as drainage area and channel slope. Presented here are algorithms and results of a new and CPU-intensive, L-moment-based approach using R and package lmomco. The approach simultaneously fits seven distributions (GEV, GLO, GNO, GPA, Kappa/GLD, LP3, PE3) at 645 stations with an aggregate record of 16,300 years. Station-specific (645) flood distributions were computed by combining the seven into one. A regional Wakeby distribution, by simulation, provides L-scale-based sampling errors, and L-scale of the seven distributions provides model errors. Combined, the errors provide WLS weights. Based on diagnostics, the new equations are reliable.

Improving Probability-Weighted Moment Methods for the Generalized Extreme Value Distribution

✱Armelle Guillou, Université de Strasbourg 1, IRMA, 7 rue René Descartes, Strasbourg cedex, 67084 France, *guillou@math.u-strasbg.fr*; Jean Diebolt, Université de Marne-La-Vallée; Philippe Naveau, Laboratoire des Sciences du Climat et de l'Environnement, IPSL-CNRS; Pierre Ribereau, Université de Montpellier

Key Words: Empirical processes, maximum likelihood estimators

In 1985,Hosking et al. estimated with the probability-Weighted Moments (PWM) method the parameters of the Generalized Extreme Value (GEV) distribution. Their approach is very popular in hydrology and climatology because of its conceptual simplicity and its good performance. Its main drawback resides in its limitations when applied to strong heavy-tailed densities. Whenever the GEV shape parameter is larger than 0.5, the asymptotic properties of the PWMs cannot be derived. To broaden the validity domain of the PWM approach, we take advantage of a recent extension of PWM to a larger class of moments. This allows us to derive the asymptotic properties of our estimators for larger values of the shape parameter. The performance of our approach is illustrated by studying simulations of small, medium and large GEV samples. Comparisons with other GEV estimation techniques are performed.

Multivariate Extension of L-Moments via L-Comoments and Applications

✱Peng Xiao, East Carolina University, East Fifth St, Greenville, NC 27858, *xiaop@ecu.edu*; Robert Serfling, University of Texas at Dallas

Key Words: L-Moments, Multivariate, Nonparametric, Coskewness, Cokurtosis, Regional frequency analysis

Multivariate statistical analysis typically relies on moment assumptions of second order and higher. With increasing interest in heavy tailed distributions, however, it is desirable to describe dispersion, correlation, skewness, kurtosis, etc., under merely first order moment assumptions. One approach consists of multivariate extension of univariate L-moments via "L-comoments," analogues of classical central moment notions of covariance, correlation, coskewness, cokurtosis, etc., but all defined under just first moment assumptions. Key properties of L-comoments and their sample versions will be reviewed. Results on L-correlational analysis will be presented. And the developing role of L-comoments as a new multivariate modeling and analysis tool in applications such as actuarial science, financial risk analysis, and environmental science will be highlighted.

388 Where Training Diverges: Nontraditional Paths to Statistical Consulting ●▲

Section on Statistical Consulting

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Application of Statistics and Measurement in the Social Sciences

✱Dale Glaser, Glaser Consulting/SDSU/AIU/USD, 3115 4th Avenue, San Diego, CA 92103, *glaserconsult@sbcglobal.net*

Key Words: Statistical consulting, Statistics and psychology

Though the training of statistics is generally associated with mathematical or statistical departments, the application of statistical methods is incorporated in a wide array of disciplines, ranging from biology to sociology to economics to education. The psychological sciences have a rich history of applying statistical techniques, rendered more complex with the type of "latent" constructs we often encounter (e.g., anxiety, depression, quality of life). This presentation will focus on how statistics are used in a consulting framework in the psychological sciences, including psychometric testing (e.g., reliability, validity), research design (e.g., experimental and quasi-experimental designs), and survey research. Moreover, the techniques that are often employed in such a context (e.g., structural equation modeling, multilevel modeling, etc.) will be detailed.

The Best of Both Worlds: Medical Science and Statistics

✱Nicole Close, EmpiriStat, Inc., 13694 Sam Hill Drive, Mount Airy, MD 21771-3900, *nclosestats@yahoo.com*

Key Words: biostatistics, clinical trials, regression, event-time data, consulting, sample size and power

This presentation will explore how the medical sciences and statistics are joined in the field of biostatistics and the challenges encountered when consulting and applying methodology for clinical trials. Challenges to be discussed include sampling methods, sample size calculations and power of a study, relative risk estimates, exact inference, large sample tests, logistic regression models and analysis of event-time data. At the conclusion, the audience will have greater knowledge of a nontraditional pathway leading to biostatistics, consulting in the field and an introduction to methodology and the challenges encountered in the field.

Psychological Scientist and Statistical Consultant: Developing and Maintaining a Dual-Career Identity

✱Todd Bodner, Portland State University/Elenchus, Inc., *tbodner@pdx.edu*

Key Words: Statistical consulting, Psychology, Statistics education, Nontraditional training

This paper recounts the experiences and challenges the author encountered in the development of a dual-career identity as a psychological scientist employed in a university setting and as a statistical consultant with academic and nonacademic clients. The discussed lessons learned may be helpful to individuals interested in a statistical consulting career who received methodological training outside of a formal degree program in statistics. Particular attention is given to how the author 1) parlayed formative experiences in his doctoral training in psychology into statistical consulting opportunities in

graduate school and beyond, 2) capitalized upon his training as a psychological scientist to bring added value to his consulting clients both in and outside of psychology, and 3) found balance between the career demands inherent in academia and statistical consulting.

Rounding Out the Numbers: Mixed Methods in Public Health Investigations

✱Lillian S. Lin, Centers for Disease Control and Prevention, 1600 Clifton Road, NE, E48, Atlanta, GA 30333, *LEL5@CDC.GOV*; Kimberlee Elmore, Centers for Disease Control and Prevention; Jennifer Galbraith, Centers for Disease Control and Prevention; Thomas M. Painter, Centers for Disease Control and Prevention

Key Words: Mixed Methods, Qualitative Methods, Public Health

Statisticians' strength as investigators is in quantitative methods and analysis. Qualitative methods complement statistical methods by elucidating how and what questions to pose, and how circumstances, relationships, social norms, and values influence findings. Social science investigations frequently employ mixed methods, using qualitative and quantitative methods in concert. Investigative methods can and must be carefully tailored to fit the study objectives. Examples of CDC public health collaborations between statisticians, who rely on quantitative methods and analysis, and social scientists, who seek to both inform quantitative measures and clarify the factors that underlie quantitative findings, will be highlighted.

Statistical Consulting in Natural Resource Management

✱Albert N. Hendrix, R2 Resource Consultants, Inc., 15250 NE 95th Street, Redmond, WA 98052, *nhendrix@r2usa.com*

Key Words: natural resource, consulting, population model, fisheries, decision analysis

Natural resource agencies such as the National Park Service and commercial entities such as hydro-electric facilities require help in making decisions under considerable uncertainty. Statistical tools that may aid decisions regarding endangered fish or mammals are biometry and population modeling (non-linear state-space models). This type of statistical modeling is typically placed into a decision analysis framework, since there are a set of management actions being considered (e.g., tear down a dam, change fishing regulations, etc.). This presentation will focus on how statistics are applied in a consulting framework in the field of natural resource sciences. Specific examples of several applications including crayfish response to Florida Everglades restoration, salmon recovery in central California, and humpback - cruise ship interactions in Glacier Bay, Alaska will be presented.

389 Counting the Dead in Iraq ●▲

Section on Statistics in Epidemiology, Section on Nonparametric Statistics, Section on Statisticians in Defense and National Security, Section on Survey Research Methods, Social Statistics Section, Scientific and Public Affairs Advisory Committee, Biometrics Section

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Iraq Data on Mortality: What Can We Believe?

✱Safaa R. Amer, National Opinion Research Center, 503 Roosevelt Blvd, Apt A509, Falls Church, VA 22044, *amer-safaa@norc.org*

Key Words: Iraq, Mortality, Survey

Estimating mortality in Iraq is unique in many ways. With the history of wars in the country, an authoritarian regime lasting for decades, the invasion, and the political winds, researchers face many challenges. Several studies have attempted to estimate mortality in Iraq for many reasons and the numbers could be as powerful as bombs taking the lives of many innocents! In this paper, we walk through the different stages of the survey cycle and outline the challenges associated with the task of estimating mortality. At the end of the paper, the reader would have to decide what to believe!

Ethical and Data-Integrity Problems in the Second Lancet Survey of Mortality in Iraq

✱Michael Spagat, Royal Holloway, University of London, Department of Economics, Egham, Surrey, International TW20 0EX United Kingdom, *m.spagat@rhul.ac.uk*

Key Words: Survey, Iraq, Falsification, Mortality, Fabrication, Ehtics

I consider the second Lancet survey of mortality in Iraq published in 2006. I give evidence of ethical violations against the survey's respondents including endangerment, privacy breeches and shortcomings in obtaining informed consent. Violations to minimal disclosure standards include non-disclosure of the survey's questionnaire, data-entry form, data matching anonymized interviewer IDs with households and sample design. I present evidence suggesting data fabrication and falsification that falls into nine broad categories.

Methods for Measuring Mortality in Conflicts: A Comparative Analyses of D.R. Congo, Iraq, and Darfur

✱Debarati Guha-Sapir, University of Louvain (CRED), 3094 Clos Chapelle aux Champs, Brussels, International 1200 Belgium, *debby.sapir@uclouvain.be*; Olivier Degomme, University of Louvain

Key Words: mortality, civil conflict, Iraq, Darfur, Congo

In recent years, mortality in civil conflicts have received increasing attention. Studies in four high visibility conflict theatres have used a variety of methods to calculate violent and non-violent deaths, often generating widely varying estimates. We review the methodological strengths and weaknesses of these methods and suggest directions for future studies of mortality in war conditions.

Nonviolence-Related Mortality in Iraq: Evidence from Iraq Family Health Survey (IFHS)

✱Mohamed M. Ali, WHO-MHI, Geneva, 1211 Switzerland, *alim@who.int*

Key Words: iraq, cause-specific mortality, excess death, survey, Uncertainty range

The recent paper "Violence-Related Mortality in Iraq, 2002–2006," published in the *New England Journal of Medicine* by the IFHS study group, focused mainly on the violence-related mortality and reported only on the unadjusted pre- and post-invasion nonviolence rates. Estimating the nonviolence deaths and the overall and cause specific excess deaths are equally important. In this paper we will analyze the IFHS mortality data with more emphasis on the nonviolence deaths and we will attempt to assess the post-invasion excess deaths. As for the violence-related deaths, adjustment methods will be used to attempt to account for under-reporting and other sources of bias in estimates of nonviolence deaths. Mortality rates by selected background characteristics will be presented and the limitations of the data and the methods used will be discussed.

390 Collaborative Projects in Statistics Education Research ▲

Section on Statistical Education

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Collaborative Projects in Statistics Education Research

✱Sterling Hilton, Brigham Young University, Department of Statistics, Provo, UT 84602, *hiltons@byu.edu*; ✱Jennifer Kaplan, Michigan State University, Department of Statistics and Probability, Michigan State University, East Lansing, MI 48823, *kaplan@stt.msu.edu*; ✱Tisha Hooks, Winona State University, Department of Mathematics & Statistics, Winona State University, Winona, MN 55987, *thooks@winona.edu*; ✱Leigh Harrell, Virginia Polytechnic Institute and State University, Department of Statistics, Virginia Tech, Blacksburg, VA 24061, *leighh@vt.edu*; ✱Diane Fisher, University of Louisiana at Lafayette, Mathematics Department, University of Louisiana at Lafayette, Lafayette, LA 70504-1010, *dfisher@louisiana.edu*

Key Words: Collaboration, Mentoring, Statistics Education, CAUSE

The field of statistics education is growing, but needs more quality studies to address unanswered questions. To meet this need, the Research Advisory Board (RAB) of CAUSE initiated a mentoring program under the NSF funded CAUSEmos grant. The program's purpose is to support the development of researchers new to statistics education. Three teams were formed from eleven faculty participants based on common research interests. During the first year, each team conducted a literature review, refined research questions, and designed a preliminary research project. In this session an RAB member will present an overview of the mentoring program. Representatives from each group will then respond to a set of questions to address participation in the mentoring program, important ideas learned from the literature, connections between the literature and proposed research, and preliminary findings.

391 Applications in Health Sciences ●

ENAR, Biometrics Section

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Impact of Inspection Errors on Average Run Length of CUSUM Charts

✱Yougui Wu, University of South Florida, 13201 Bruce B Downs Blvd, Tampa, FL 33612, *ywu@health.usf.edu*; Yiliang Wu, University of South Florida; Wang Wei, University of South Florida

Key Words: inspection errors, control chart, Average Run Length, CUSUM

The cumulative sum (CUSUM) chart is commonly used for monitoring the counts of defective items which are usually well modelled by a binomial distribution with parameters n and p where n is the number of items inspected each time and p is the process fraction of defective items produced. Conclusive evidence exisits to show that attributes inspective tasks are far from error free. Even under ideal inspection conditions,error rates exceeding 25 percent are not uncommon.The purpose of this paper is to determine and illustrate the effect of inspection error on average run length of CUSUM charts. We show that the effect can be quantified by the square of the correlation coefficient between the correct and the incorrect inspections. In addition, a procedure for determining the sampling size that compensates for inspection error will be developed and illustrated.

Bayesian Distributed Lag Models

✱Mark J. Meyer, American University, 4400 Massachusetts Ave NW, Department of Mathematics and Statistics, Gray Hall 103, Washington, DC 20016, *mark.john.meyer@gmail.com*; Elizabeth J. Malloy, American University; Sara D. Adar, University of Washington; Brent A. Coull, Harvard School of Public Health

Key Words: Moving Average, Distributed Lag Model, Bayesian methods, Panel Study, Monte Carlo Simulation, Mixed Model

Studies examining the acute health effects of pollution exposure frequently involve covariate effects measured discretely over time. Moving average and distributed lag models are two commonly used methods for incorporating time-varying covariates into linear models. Moving averages require fitting sequential models, with each including an exposure average taken successively further back in time. Distributed lag models use lagged exposures as covariates. In a simulation study, we examine properties of moving average models and two distributed lag models: a parametric and a penalized. We structure the simulations on panel studies of the acute health effects of pollution exposure with repeated exposure measures. The moving average and parametric distributed lag models are fit using classical mixed model methods. We use a Bayesian distributed lag model for examining the penalized model.

Estimating Benefits Due to Fecal Occult Blood Test in Colorectal Cancer Screening

✱Dongfeng Wu, University of Louisville, 555 S. Floyd Street, Suite 4026, Louisville, KY 40202, *dongfeng.wu@louisville.edu*; Gary L. Rosner, University Texas MD Anderson Cancer Center

Key Words: sensitivity, lead time, sojourn time, cancer screening, transition probability, colorectal cancer

We applied the statistical methods developed by Wu, Rosner and Broemeling 2005, 2007 using the Minnesota colorectal cancer data, to make Bayesian inference for the age-dependent sensitivity, transition probability, sojourn time distribution, and lead time distribution, for both genders in a periodic screening program. The sensitivity appears to increase with age for both genders. However, the posterior mean sensitivity is not monotonic for males; it has a peak around age 74. The age-dependent transition probability is not a monotone function of age; it has a single maximum at age 72 for males and a single maximum at age 75 for females. The posterior mean sojourn time is 4.08 years for males and 2.41 years for females. To guarantee a 90% chance of early detection, it seems necessary for the males to take the test every 9 months, while the females only need to take it annually.

A Gene Selection Method for GeneChip Array Data with Small Sample Sizes

✱Zhongxue Chen, The Children's Hospital of Philadelphia, Biostatistics and Data Management Core, Children's Hospital of Philadelphia, philadelphia, PA 19104, *chenz@email.chop.edu*

Key Words: Multiple comparison, False discovery rate, Microarray, gene selection

Multiple comparison issue is a challenge in gene selection of microarray data. Current methods that control FDR are not applicable when sample sizes are very small due to the poor estimations for p-values and the distribution of true null. We propose a model-based information sharing method, which takes the advantage of information shared among genes and has better test statistics. After appropriate data transformation, we use a normal distribution to model the mean differences of true nulls. The parameter in the model is estimated by all data. The p-values are calculated from the model. Instead of controlling FDR,we select genes for a given cutoff p-value and then estimate the false discover rate. We compare our new method with others by simulation and real microarray data. The comparison results show that this new method is more reliable.

Telescoping Determinant Plots as an Index of Cognitive Information Accounted for in Electrophysiological Visual Evoked Response Data

✱Daniel S. Barron, Brigham Young University, 1001 SWKT, Provo, UT 84602, *dansbarron@hotmail.com*; Bruce L. Brown, Brigham Young University; Dawson W. Hedges, Brigham Young University

Key Words: visual evoked response (VER), electrophysiology, cognitive processes, multivariate graphics

Brown, Hedges, and Gantt (2008) introduce a multivariate graphical technique for capturing cognitive processes through visual evoked responses (VER). In the past it has been difficult to capture cognitive information in VER profiles because, like fingerprints, they are strongly ideographic. That is, they have a strong individual difference component, with each person having a unique set of VER profiles. However, aperiodic components analysis has the demonstrated capacity to separate the surface structure of individual differences from the deep structure of underlying cognitive processes, both mathematically and graphically. This paper demonstrates a secondary development in aperiodic components methodology—a procedure for quantifying the extent to which each aperiodic component captures a given aspect of cognitive processing. This procedure is based upon telescoping determinant plots.

392 ROC and Diagnostic Methods ●

Biometrics Section, Section on Nonparametric Statistics

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Nonparametric Estimation for Time-Dependent AUC

✱Hung Hung, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan, *double.red@gmail.com*

Key Words: AUC, bivariate estimation, Gaussian process, nonparametric, ROC, survival data

The area under the receiver operating characteristic (ROC) curve (AUC) is one of the commonly used measure to evaluate or compare the predictive ability of markers to the disease status. Motivated by an angiographic coronary artery disease (CAD) study, our objective is mainly to evaluate and compare the performance of different biomarkers in the prediction of CAD-related vital status over time. Under the random censoring and conditionally independent censoring mechanisms, a class of nonparametric estimators is proposed for the time-dependent AUC. The algebraic and geometric properties of the estimators are also provided in this article. Meanwhile, the asymptotic Gaussian processes of the estimators are established to construct the approximated confidence regions of the AUCs and test for the equality of AUCs of plasma biomarkers.

Tree-Structured Analysis for Determining Optimal Diagnostic Tests for Patients

✱Caixia Li, University of California, San Francisco, Dept of Radiology Box0946, 185 Berry Street Suite 350, San Francisco, CA 94143-0946, *caixia.li@radiology.ucsf.edu*; Ying Lu, University of California, San Francisco

Key Words: Tree-structured algorithm, Cost effectiveness, Subgroup analysis, Diagnostic test

This paper presents a new tree-structured algorithm to determine "who should take which test" according to risk factors. Our model chooses from diagnostic tests 1 and 2. A positive test will be followed with a treatment. The algorithm assigns subjects into one of the four action groups: (0) subjects with low risk such that no need of diagnosis and treatment; (1) subjects for whom test 1 is more cost-effective than test 2; (2) subjects for whom test 2 is more cost-effective than test 1; (3) subjects have high enough disease risk who should be treated without any test. The splitting criterion is to maximize the cost-effective difference, which is the difference between incremental cost and quality adjusted life year gain (in $). Several pruning criteria are explored. The proposed method is applied to determine who should receive DXA versus QUS to determine osteoporosis treatment.

Using Marginal ANOVA Models To Motivate, Generalize, and Derive Properties for the Obuchowski-Rockette Procedure for Multireader ROC Data Analysis

✱Stephen Hillis, VA Iowa City Medical Center, CRIISP (152), 601 Highway 6 West, Iowa City, IA 52246-2208, *steve-hillis@uiowa.edu*

Key Words: marginal ANOVA model, receiver operating characteristic (ROC) curve, Dorfman-Berbaum-Metz (DBM), Obuchowski-Rockette, diagnostic tests, diagnostic radiology

Although the Dorfman-Berbaum-Metz method has been a popular method for analyzing multireader ROC studies, statistical properties are difficult to derive since it is based on a model with pseudovalues as outcomes. Recently it has been shown that equivalent results can be obtained using the Obuchowski-Rockette (OR) ANOVA model with correlated errors. However, a drawback of the OR model is that it is not a familiar model and properties are tedious to derive. In this talk I show how the OR model can be viewed as a marginal ANOVA model resulting from a conventional ANOVA model with independent errors. Viewing the OR model in this way provides an intuitive motivation for the OR model and generalizations of it, an easy way to derive the appropriate OR test statistics and its properties, and a unifying framework that clearly shows the relationship between the DBM and OR methods.

Time-Dependent Receiver Operating Characteristic (ROC) Curves for Design of Early Alzheimer's Disease Trials

✱Norberto Pantoja-Galicia, Harvard University, Department of Biostatistics, 655 Huntington Avenue. Building 2, 4th Floor, Boston, MA 02115, *npantoja@hsph.harvard.edu*; Rebecca Betensky, Harvard University

Key Words: Intermediate endpoint, Prediction, Receiver operating characteristic curve

Clinical trials for individuals who are at the earliest stages of Alzheimer's disease (AD) are of interest, as there is promise of effective treatment for these individuals. However, these trials cannot use progression to AD as their endpoint due to the long time course of this transition. Thus, we use longitudinal data to find intermediate endpoints (IE) that will be observed with reasonable frequency in a typical follow-up period of a clinical trial that have an optimal prediction profile for progression from pre-Mild Cognitive Impairment to AD. For this purpose, we estimate time dependent ROC curves that take into account the censored nature of the data. We investigate methods for the comparison of the associated sequences of areas under these curves for different candidate IE, with the goal of identifying an optimal IE for use as the endpoint in future trials.

Estimation for the Optimal Combination of Markers Without Modeling the Censoring Distribution

✱Chin-Tsang Chiang, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan, *chiangct@ntu.edu.tw*

Key Words: AUC, IPW, optimal composite marker, ROC curve, survival data, vital status

In the time-dependent receiver operating characteristic curve (ROC) analysis with several baseline markers, research interest focuses on seeking appropriate composite markers to enhance the accuracy in predicting the vital status of individuals over time. Based on censored survival data, we proposed a more flexible estimation procedure for the optimal combination of markers under the validity of a time-varying coefficient generalized linear model (GLM) for the event time without restrictive assumptions on the censoring pattern. In contrast, the inverse probability weighting (IPW) approach might introduce a bias when the selection probabilities are misspecified in the estimating equations. The performance of both estimation procedures are examined and compared through a class of numerical studies.

Probability Sampling Framework for Survival Analysis with Time-Varying Covariates

✱Stephen L. Rathbun, The University of Georgia, 132B Coverdell Center, University of Georgia, Athens, GA 30602, *rathbun@uga.edu*

Key Words: Ecological Momentary Assessment, Smoking Lapse, Time-Varying Covariates, Survival Analysis

Current approaches to survival analysis with time-varying covariates involve the joint modeling of the time-varying covariates and the lifetimes to events. Model misspecification can lead to biased estimates of survival model parameters. I introduce a new approach based on probability-sampling of the time-varying covariates during the lifetimes of the events. A design-unbiased estimator of the cumulative hazard is substituted into the log likelihood which is then maximized to obtain the proposed parameter estimator. This estimator is consistent and asymptotically normally distributed. Simulations show that the proposed estimator competitive with the conditional score estimator when the model is correctly specified, and far superior when it is misspecified. This approach is illustrated on the lifetime to smoking lapse from an Ecological Momentary Assessment of smoking.

A Bayesian Framework for Assessing Longitudinal Agreement Between Two Continuous-Scale Diagnostic Tests in the Presence of Established Thresholds

✱Farzad Noubary, Harvard School of Public Health, *noubary@fas.harvard.edu*; Michael Hughes, Harvard School of Public Health

Key Words: longitudinal, stochastic, Bayesian, HIV, CD4, resource-poor

The World Health Organization (WHO) has urged the development of inexpensive CD4 methodologies to guide clinical decisions regarding initiation and monitoring of therapy for HIV-infected patients in resource-poor settings. Most clinical papers studying agreement between newly developed CD4 assays and gold standard methods have simply reported cross-sectional results such as sensitivity and specificity relative to thresholds for initiating therapy. We use a stochastic model with an integrated Ornstein-Uhlenbeck process to study the longitudinal question of how timing of treatment initiation changes when inexpensive assays are used. We also investigate the influence of frequency of measurement and use of a confirmatory value on treatment starting times. A Bayesian framework is used to incorporate uncertainty in the estimation of the fixed, random, and stochastic parameters in the model.

393 Which Way To Weight? Bias vs. Variance

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Weight Smoothing Models in Clustered or Cross-Classed Sample Designs

✱Shelby (Xiaobi) Huang, The University of Michigan, 2013 Medford Road, Apt. H162, Ann Arbor, MI 48104, *xiaobih@umich.edu*; Marc N. Elliott, RAND Corporation

Key Words: Sampling weights, clustered sample designs, cross-classed sample designs, weight trimming, weight smoothing, variable selection

Highly disproportional sample designs have large weights, which will introduce undesirable variability in statistical estimates. Weight trimming fixes a cutpoint weight and sets larger weights to this cutpoint value, reducing variability at the cost of introducing some bias. Previous work developed Bayesian "weight smoothing" models to produce general model-based weight trimming estimators of population statistics, but has been limited to the context of stratified and post-stratified sample designs. This paper extends the "weight smoothing" methodology to a more general class of complex sample design that include single or multistage cluster samples and/or strata that "cross" the weight strata. The methods are applied to linear regression models. We also discuss to use the variable selection approach to explore more flexible and robust "weight smoothing" models.

An Empirical Comparison of Procedures for Reducing Bias Due to Weighting Cell Collapsing

✱Linda Tompkins, National Center for Health Statistics, 3311 Toledo Road, Room 3115, Hyattsville, MD 20782, *lit3@cdc.gov*; Jay J. Kim, National Center for Health Statistics

Key Words: poststratification, coverage, cell collapsing

Post-stratification is a common ratio-adjustment method for adjusting sampling weights for differential coverage in household surveys. Weighting cells are formed based on known characteristics of sample respondents for which population controls are available from the census or other sources. The inverses of post-stratification factors are usually referred to as coverage ratios, which can vary among demographic subgroups. A standard procedure in weighting to correct for this is to collapse cells for small race groups with those for Whites. It has been shown that this practice may introduce bias in the estimates for the race groups with smaller coverage ratios, say American Indians and Asians. In this paper, two procedures developed to reduce potential bias in these cells are empirically compared to the current weighting approach using the 2005 National Health Interview Survey.

Weight Development for Outliers in a Panel Sample

✱Yan K. Liu, Statistics of Income, IRS, P.O.Box 2608, Washington, DC 20013, *yan.k.liu@irs.gov*; Fritz Scheuren, The University of Chicago; Victoria Bryant, Statistics of Income, IRS

Key Words: panel sample, stratified sample design, sampling weight

Statistics of Income of IRS started a panel sample in tax year 1999. It was a stratified sample of individual returns, where stratum boundaries were based on the individual income. Because of the large number of population returns and the much skewed income distribution, sampling rates across strata were quite different. This poses a problem for 'stratum-jumper' returns that shift between strata in different years, as their yearly income change dramatically. This particularly affects returns that were selected with low sampling rates, thus being assigned higher base weights. If the income on a return grows such that it moves into a stratum with a lower probability of selection, then the associated original weight is no longer appropriate for cross-sectional estimation. In this paper, we review the ad hoc procedures and also develop more options for the treatment.

Alternative Methods To Adjust for Nonresponse in the American Community Survey (ACS)

✱Robyn Sirkis, U.S. Census Bureau, Washington, DC 20230, *robyn.b.sirkis@census.gov*

Key Words: American Community Survey, weighting, estimation, nonresponse, propensity scores

This paper discusses research on alternative methods to the ACS weighting methodology to adjust for nonresponse, which remains unchanged since 1996. The study universe is the 2006 ACS tabulation sample. Methods using the inverse of the propensity score as the adjustment factor, the mean inverse of the propensity score as the adjustment factor, and forming the adjustment cells using the propensity scores were investigated. The methods were tested on a data set where a sample of respondents were recoded to nonrespondents and compared to the original data set. The data set that was recoded was compared to the data set that was not recoded by analyzing several population, household, and housing unit estimates. The variances were analyzed to ensure there is not an increase due to the alternative methodologies.

Evaluating Alternative Use of Population Controls for American Community Survey Weighting Methodology

✱Mark E. Asiala, U.S. Census Bureau, DSSD HQ-4H057, 4600 SILVER HILL RD, Washington, DC 20233, *mark.e.asiala@census.gov*; Michael Beaghen, U.S. Census Bureau; Keith Albright, U.S. Census Bureau

Key Words: American Community Survey, weighting, population estimates, bias, variance

The objective of this research is to assess whether the use of independent population estimates produced by the Census Bureau's Population Estimates Program improves the estimates produced by the American Community Survey (ACS). Six alternatives to the current use of population controls were investigated by applying the controls either at higher levels of geography or with less demographic detail. An expected worst-case bias scenario for the population estimates is studied by applying the April 1, 2000 population estimates based on the 1990 Census as controls to the Census 2000 Supplementary Survey data. Various county-level demographic estimates were calculated for each of the seven methods. The county-level relative bias, variances, and relative mean square errors to the Census 2000 data were then calculated to compare the relative quality of the estimates for each method.

Estimation for a Multipurpose and Multiframe Survey: The National Immunization Survey - Adult

✱Hee-Choon Shin, National Opinion Research Center, 2232 University Dr., Naperville, IL 60565, *shinh@uchicago.edu*; Noelle-Angelique Molinari, NCIRD/CDC; Meena Khare, National Center for Health Statistics; Kirk Wolter, National Opinion Research Center

Key Words: Muitiframe Design, Dualframe Design, Vaccination Coverage, Pseudo-Maximum Likelihood Estimator

In a multiframe survey, samples are drawn independently from overlapping frames that are assumed to cover the population of interest. The National Immunization Survey-Adult (NIS-Adult) is a national telephone survey of adults aged 18 years and older collecting data on vaccination-related variables. The NIS-Adult selected samples from two frames: the National Health Interview Survey (NHIS) and the age-targeted list of phone numbers. The estimators of Hartley and of Fuller and Burmeister would be efficient estimators for a single response variable. In reality, there are many response variables which require separate weights for efficient estimators. An alternative (pseudo-maximum likelihood estimator, PMLE) which utilizes the same survey weights for all response variables was proposed by Skinner and Rao. The three alternative estimators will be developed for the NIS-Adult.

Analyzing Weighting Methods in the Federal Human Capital Survey

✱Taylor H. Lewis, U.S. Office of Personnel Management, 1900 E St., NW, Room 7439, Washington, DC 20415, *Taylor.Lewis@opm.gov*; Katie Joseph, U.S. Office of Personnel Management; Eulus Moore, U.S. Office of Personnel Management

Key Words: survey, nonresponse, weighting, raking

This paper compares two weighting procedures for the Federal Human Capital Survey, a sample of nearly 440,000 federal employees stratified by agency and supervisory status to gauge perceptions of various workplace issues. The first weighting method involves a basic nonresponse adjustment using stratum information only. The second method utilizes extra demographic data on the frame known of both respondents and nonrespondents in a CHAID analysis to construct nonresponse cells, concluding with a raking step. This paper interprets results from a nonrespondent follow-up to explain possible mechanisms of nonresponse and compares means and variances of estimates in an effort to evaluate which method is more appropriate.

394 Alternative Measures of Association ●

Section on Statistics in Epidemiology, Biometrics Section

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

On Effect-Measure Modification: Relationships Among Changes in the Relative Risk, Odds Ratio, and Risk Difference

✱Babette A. Brumback, University of Florida, Department of Epidemiology and Biostatistics, College of Public Health and Health Professions, Gainesville, FL 32610, *brumback@phhp.ufl.edu*; Arthur Berg, University of Florida

Key Words: heterogeneity, common outcomes, clinical trials, cohort studies, cross-sectional studies, case-control studies

It is well known that presence or absence of effect-measure modification depends upon the chosen measure. What is perhaps more disconcerting is that a positive change in one measure may be accompanied by a negative change in another. Therefore, research demonstrating that an effect is 'stronger' in one population when compared to another, but based on only one measure, for example the odds ratio, may be difficult to interpret for researchers interested in another measure. This talk reports on an investigation of relations among changes in the relative risk, odds ratio, and risk difference from one stratum to another. Analytic and simulated results are presented concerning conditions under which the measures can and cannot change in opposite directions. Data-analytic and hypothetical examples are also presented.

A Generalized Estimator of the Attributable Benefit of an Optimal Treatment Regime

✱Jason Brinkley, North Carolina State University, Department of Statistics, 27695-8203, *jsbrinkl@ncsu.edu*; Anastasios Tsiatis, North Carolina State University; Kevin Anstrom, Duke Clinical Research Institute

Key Words: Optimal Treatment Regime, Causal Inference, Attributable Risk, Attributable Benefit

For many diseases where there are several treatment options it is often the case that there is no consensus on the best treatment to give individual patients. If a treatment strategy is defined as an algorithm which dictates which treatment patients receive based on factors that are measured on that individual, then the ultimate goal is to identify the optimal treatment strategy which will result in the smallest proportion of poor outcomes. As a measure of the utility of a particular treatment strategy we propose using the proportion of events that could have been prevented had the optimal treatment strategy been used. Using ideas from causal inference we will formally define the utility measure of interest and derive estimators for this measure and the large sample properties of these estimators. Finally an example will be put forth to illustrate the usefulness of this measure.

Estimating Model-Adjusted Risk and Prevalence Ratios from Survey Data in SUDAAN Release 10

Gayle Bieler, RTI International; ✱Rick L. Williams, RTI International, 2816 Stone Gap Ct, Raleigh, NC 27612, *rwilliams@rti.org*; George G. Brown, RTI International

Key Words: risk ratios, prevalence ratios, logistic regression, survey data analysis, conditional and predicted marginals

There is increasing interest in estimating and drawing inferences about risk or prevalence ratios and differences instead of odds ratios in the regression setting (Greenland, Am J Epid, 2004). Recent publications have shown how SAS PROC GENMOD is used to estimate these parameters in non-population based studies (Spiegelman, et al, Am J Epid, 2005). This paper shows how adjusted risks, risk differences, and risk ratio estimates can be obtained from multinomial logistic regression models in the complex sample survey setting, which yield population-based inferences. SUDAAN 10 will be used to obtain point estimates, standard errors (via robust variance methods), and confidence limits for the parameters of interest. Adjusted risks and risk ratios are obtained from conditional or predicted marginals in the fitted regression model. The 2006 NHIS data will be used to illustrate these concepts.

Evaluation of Population Attributable Risk in the Presence of Competing Causes

✱Jie Huang, Northwestern University, 680 N. Lakeshore Dr. Suite 1102, Chicago, IL 60611, *jjhuang@northwestern.edu*

Key Words: population attributable risk, competing risks, propensity score

Population attributable risk is an important measure used in epidemiology. However, in the presence of competing risk, the population attributable risk with respect to a single cause could provide misleading information. We propose to simultaneously estimate the profile of the population attributable risk for all causes using the propensity score approach. The efficiency enhancement methods will be employed. The proposed methods will be illustrated with a practical example as well as numerical studies.

Changing the Reference Group in Published Odds Ratios for Multilevel Categorized Exposures in Preparation for Meta-Analysis

✱Monika M. Wahi, Byrd Alzheimer's Institute, 4001 E. Fletcher Ave., 2nd floor, FAST Dept, Tampa, FL 33613, *mwahi@byrdinstitute.org*; Skai W. Schwartz, USF COPH

Key Words: Statistics, Meta-analysis, Odds Ratio, Methods, Epidemiology

Challenges to conducting a meta-analysis include combining information (1) from analyses with variable levels of adjustment; (2) using incomparable populations across studies; (3) using different categorizations of a continuous variable and/or (4) different reference groups. This paper explores standardizing the reference group across multiple studies. For two levels of exposure, this is a straightforward process. For example, if the measure of association is an odds ratio (OR), one simply takes the reciprocal of the OR and the lower and upper bounds of the confidence interval. However, with a multi-level exposure, obtaining new confidence intervals is hampered by lack of a covariance estimate. Care should be taken to report ORs using reference groups consistent with those used in previous articles.

Generalized Linear Mixed Models in Oral Health Research

✱Ileana Baldi, University of Torino, Cancer Epidemiology Unit, Via Santena 7, Turin, International 10126 Italy, *ileana.baldi@cpo.it*; Dario Gregori, University of Torino

Key Words: Generalized Linear Mixed Model, Linear Mixed Effect Model, Heteroscedasticity, Gamma

In case of positive and skewed data, a common approach is to fit a linear model to the log-transformed response, with the parameters being eventually evaluated after a back-transformation on the original scale. This method is known to be biased, with the bias increasing with the heterogeneity in data. An alternative approach based on the Generalized Linear Mixed Model (GLMM) is proposed. We evaluate the performance of the Gamma GLMM under different data generating mechanisms and compare it to that of the Linear Mixed Effect Model on a log-transformed response. A case study from oral health epidemiology is analyzed and discussed. The simulation study shows that the method of fitting linear models to a log-transformed response may have relatively little bias if the gamma scale parameter is constant, but suffers from substantial bias if the scale parameter varies with the covariates.

On the Confidence Interval Estimates of Impact Numbers for Cross-Sectional Sampling Schemes

✱Tanweer Shapla, Eastern Michigan University, 515 Pray-Harrold, Department of Mathematics, Ypsilanti, MI 48197, *tshapla@gmail.com*

Key Words: Attributable risk, Impact numbers, Confidence interval, Cross-sectional sampling

While relative risk difference and attributable risk are frequently used effect measures in epidemiological researches designed to describe the effect of exposures on disease development, recently proposed impact measures such as EIN, CIN, and ECIN have become popular because of their interpretational simplicities. This paper concentrates on the development of the confidence interval construction for these impact numbers under cross-sectional sampling scheme. A real life example on serum cholesterol level and coronary heart disease is used to compute and interpret these impact numbers.

395 Association of Genetic Factors with Outcome ●

Section on Statistics in Epidemiology, Biopharmaceutical Section, Biometrics Section
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Analysis of Population-Based Genetic Association Studies Using Propensity Scores

✱Huaqing Zhao, University of Pennsylvania, The Children's Hospital of Philadelphia, BDMC, 3535 Market Street, Suite 1035, Philadelphia, PA 19104, *zhao@email.chop.edu*; Nandita Mitra, University of Pennsylvania; Timothy R. Rebbeck, University of Pennsylvania

Key Words: population-based genetic association, propensity score, confounding, population stratification, admixture, prostate cancer

Propensity scores are widely used to address confounding in observational studies. However, they have not been adopted to deal with population stratification in genetic association studies. Population stratification is a bias that arises when differences in allele and disease frequencies exist. Existing methods (genomic control, structured association) control this bias using genetic markers alone. However, confounding can arise due to non-genetic factors as well. We propose a propensity score approach that controls for both genetic (null and admixture informative markers) and non-genetic covariates. We illustrate this approach in a case-control study of CYP3A genotypes and prostate cancer. We compare the proposed method with existing methods using simulations. We conclude that our method provides a novel and robust tool for obtaining less biased estimates of genetic associations.

Application of Statistical Methods in Detecting Gene-Gene Interactions for a Cancer Case-Control Study

✱Hui Zhao, The University of Texas MD Anderson Cancer Center, 1155 Herman P.Pressler DR., Unit 1365, Houston, TX 77030, *huizhao@mdanderson.org*; Qingyi Wei, The University of Texas MD Anderson Cancer Center

Key Words: SNPs, gene-gene interaction, multifactor dimensionality reduction, Tukey's one-degree-of-freedom model, classification tree modeling

Interactions between genes and environmental factors are known to play a role in the etiology of cancer. Since the conventional logistic regression model lacks power for detecting gene-gene interactions when multiple genes are involved, various data mining and machine learning techniques have been developed to analyze these interactions for an association study. In the current study, we applied four existing methods-multifactor dimensionality reduction, Tukey's one-degree-of-freedom model, the classification tree model, and the conventional logistic regression model-to detect gene-gene interactions. We first evaluated these methods using simulated genetic data and then applied them to a cancer case-control dataset. We also discuss some ideas about how these methods may help us understand the association between genetics and cancer risk.

Relationship Between Chronic Obstructive Pulmonary Disease and TLR4 Polymorphism (Asp299Gly)-BMI Interaction in a Canadian Population

✱Punam Pahwa, University of Saskatchewan, 103 Hospital Drive, Dept. of Community Health and Epidemiology, Saskatoon, SK S7N OW8 Canada, *pup165@mail.usask.ca*; Donna Rennie, University of Saskatchewan; Chandima P. Karunanayake, University of Saskatchewan; James Dosman, University of Saskatchewan; Claire F. McGuigan, University of Saskatchewan

Key Words: Non-linear, Genetic, longitudinal, epidemiology, TLR4, Polymorphism

There is limited information concerning the statistical modeling of longitudinal data in genetic epidemiological studies with respect to gene-environmental interactions. The objective of this report is to extend existing advanced statistical approaches to the analysis of available genetic and longitudinal environmental data. We determined the effects of genetic and environmental factors and their interactions on longitudinal changes in the prevalence of respiratory symptoms in the population of Humboldt, Saskatchewan, Canada. Non-linear multivariable regression analysis based on quasi-likelihood was utilized to investigate the influence of TLR4 Asp299Gly polymorphism and body mass index interaction (adjusting for other potentially important covariates, confounders and interactions) on the longitudinal prevalence of chronic obstructive pulmonary disease.

The Effects of Genetic and Environmental Factors on Longitudinal Changes in Pulmonary Function Values in a Rural Population: The Humboldt Study

✱Chandima P. Karunanayake, University of Saskatchewan, 103 Hospital drive, CCHSA, Saskatoon, SK S7N 0W8 Canada, *cpk646@mail.usask.ca*; Punam Pahwa, University of Saskatchewan; Donna Rennie, University of Saskatchewan; James Dosman, University of Saskatchewan; Claire F. McGuigan, University of Saskatchewan

Key Words: linear mixed effects models, restricted maximum likelihood, pulmonary function values, TLR4 A299G, gene - environment interaction

The interaction between gene and environment for lung health has recently come under investigation and is not well described. The objective of this study was to determine the effects of genetic and environmental factors and their interactions on longitudinal changes in pulmonary function values. We examined lung health and associated genetic factors in two cross sectional lung studies conducted in 1993 and 2003 in the rural community of Humboldt, Saskatchewan, Canada. There were 769 subjects who participated at both time points. Information included: demographic factors, respiratory health indicators, history of allergy, smoking history, home environment, farm environment, and pulmonary function. Genetic polymorphic data for TLR4A299G was collected in 2003. The longitudinal data analysis was conducted using linear mixed effects models based on restricted maximum likelihood approach.

Application of Bayesian Classification with Singular Value Decomposition Method in a Genome-Wide Association Study

✱Soonil Kwon, Cedars-Sinai Medical Center, Medical Genetics Institute, 8700 Beverly Blvd. 665W, Los Angeles, CA 90048, *soonil.kwon@cshs.org*; Xiuqing Guo, Cedars-Sinai Medical Center

Key Words: Bayesian Classification, Singular Value Decomposition, Genome-wide association studies, Genetic

Genome-wide association studies are used to test for association over the entire genome and usually involve several hundred thousand of markers (m). The high cost of phenotyping often limits sample size (n). In order to analyze multiple SNPs simultaneously when m>>n, we introduced the Iterative Bayesian variable selection (IBVS) and have successfully applied the method to the simulated Rheumatoid Arthritis data provided by the Genetic Analysis Workshop 15 (GAW15). However, the IBVS still requires a relatively long run-time, which limits its applications. We introduce a Bayesian classification with singular value decomposition (BCSVD) method and apply it to the same GAW15 data. The BCSVD performed as good as the IBVS with less than an hour run-time compared to several days for the IBVS. We conclude that the BCSVD is good for analyzing large scale association data when m>>n.

A New Estimate of Family Disease History Providing Improved Prediction of Disease Risks

✱Rui Feng, The University of Alabama at Birmingham, Department of Biostatistics, Birmingham, AL 35294, *rfeng@ms.soph.uab.edu*; Leslie McClure, The University of Alabama at Birmingham; Hemant K. Tiwari, The University of Alabama at Birmingham; George Howard, The University of Alabama at Birmingham

Key Words: family history, family risk, family history score, age of onset, heart disease

Complex diseases often aggregate within families and using the history of family members' disease can increase the accuracy of the risk assessment and allow clinicians to better target on high risk individuals. However, available family risk scores can not reflect age of onset, gender, and family structures simultaneously. We propose a new family risk score that incorporates the age of onset and gender of family members and the family relationship. Via simulation, we demonstrate that the new score is more closely associated to the true family risk for the disease and more robust to family sizes than two existing methods. We apply our proposed method and the two existing methods to a study of heart disease. The results show that accessing family history can improve the prediction of disease risks and the new score has strongest associations with both myocardial infarction and stroke.

Selection Bias and Imputation in Genetic Association Analyses of Time-to-Event Data in Framingham Offspring Study

✱Xiaoyan Yin, Boston University, 36 Wilson Drive, Framingham Heart Study, Framingham, MA 01702, *xyin@bu.edu*; Martin G. Larson, Boston University

Key Words: bias, genetic, survival, EM, imputation

The second generation of Framingham Heart Study was genotyped partially because the DNA samples were collected about 30 years after enrollment. Participants who died before sample collection could not contribute their blood samples to the study. The missingness of genetic data is not completely at random. We studied the bias caused by analyzing the data from genotyped individuals only for genetic association with time to event outcome. We also developed easy-applicable approaches to imputing genotype data for individuals who were not genotyped. The direction of bias is away from the null hypothesis. The magnitude depends on the genetic effect size and genotype frequency. Two methods were used for data augmentation: multiple imputation and EM by the method of weights. Both methods reduce the bias dramatically than complete case analyses. The advantage and disadvantages are discussed.

396 From Teaching Business Students to Teaching Paradoxes ●▲

Section on Statistical Education

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Strategies for Increasing the Effectiveness of Statistics Teaching in Business Curricula

✱Bodapati V. Gandhi, University of Puerto Rico, Calle Azorin 710, Mansiones de Espana, Mayaguez, PR 00682, *bvrgandhi@yahoo.com*

Key Words: business statistics, teaching, textbooks, curriculum

Instructors in business statistics face two constraints: the limited time to teach their course, and the knowledge and skills students have coming into the course. Textbooks have changed considerably over the last three decades, due to changes in technology as well as the addition of cases and the introduction of additional statistical techniques. The structure of degree requirements has often not kept up pace. For students to fully integrate their knowledge of statistics as a research tool into their future course and professional work, changes are necessary to texts, prerequisite courses and related courses. The paper also discusses the efforts academic committees and administrators should make in order to concretize these changes.

A 2005–2006 Senior Fulbright Outreach to Lviv Ukraine, Improving Business Decisionmaking with Statistical Methods

✱J. Marcus Jobe, Miami University, 301 Upham Hall, Decision Sciences Dept., Oxford, OH 45056, *jobejm@muohio.edu*

Key Words: Ukraine, Senior Fulbright Scholar, Statistical Methods

The statistical lecture series/curriculum development accomplished as a 2005–2006 Senior Fulbright Scholar to Lviv Institute of Management focused on improving business decision making. In this paper, the educational activities, goals, successes and shortcomings associated with my 2005–2006 Senior Fulbright Scholar work at Lviv Institute of Management, Lviv, Ukraine are presented. Specifically, curriculum development (English/Ukrainian) and the lecture series taught to Ukrainian students on introductory business statistics, quality improvement and data mining are discussed. Aspects of the courses taught, Ukrainian higher education methodologies and cultural differences with respect to statistical understandings are presented.

Sampling from Finite Multivariate Populations

✱Bruce Barrett, The University of Alabama, Box 870226, Tuscaloosa, AL 35487-0226, *bbarrett@cba.ua.edu*

Key Words: paired data, multivariate sampling

Many students are initially exposed to the idea of a multivariate data when the topic of "paired samples" is introduced. The typical scenario is a set of measurements taken before and after some treatment. Most texts pay great tribute the reduction in variance that can be achieved when estimating the mean difference, so much so, that students often come away with the sense that paired, or multivariate measurements are always preferred over independent samples. We offer examples that help illustrate when independent samples are preferred over the matched samples. Also, we discuss how to handle issues of correlation that occur when independent samples, taken from finite populations, result in multiple measurements contributed by the same experimental unit.

Plate Expectations: A Dickens of a Problem

✱Milton W. Loyer, The Pennsylvania State University, 917 Emily Drive, Mechanicsburg, PA 17055, *loyer@lycoming.edu*

Key Words: conditional probability, paradox

A paradox is presented wherein it appears that the unconditioned probability of an event is less than the probability of the event for each of the mutually exclusive and exhaustive conditions under which it can occur. The paradox is introduced using batting averages and then presented in the real-world context in which it was first known to be observed. It will be demonstrated that no data can exhibit both this paradox and the distinctly different phenomenon known as Simpson's paradox.

Are Textbooks Fairly Priced?

✱David P. Doane, Oakland University, Dept of Decision and Information Sciences, Rochester, MI 48309-4493, *doane@oakland.edu*; Lori E. Seward, University of Colorado Denver

Key Words: textbook prices, electronic books, used books, on-line textbooks, book buyers, education costs

Many students complain that textbooks cost too much, and some states are considering legislation to force textbook prices down. But are these concerns justified? This paper reports a survey of students in a variety of business statistics classes at two universities, asking where they got the textbook, the type of textbook (new, used, international, electronic), and how much was paid. Our data set includes computer ownership, laptop use in class, perceptions of the value of textbook CDs, and reasons why students might be willing to purchase a cheaper electronic textbook. Historic data on inflation and improving real income suggest that per-page retail textbook prices may effectively be lower than 40 years ago. Our paper discusses factors affecting today's textbook market (e.g., the resale market, internet sales, qualitative changes in textbooks, and changing student lifestyles).

A Counterintuitive Regression Phenomenon

✱Aiyou Chen, Bell Labs, Alcatel-Lucent, 2c-358, 600 Mountain Ave, Murray Hill, NJ 07974, *aychen@research.bell-labs.com*; Thomas Bengtsson, Bell Labs, Alcatel-Lucent; Tin K. Ho, Bell Labs, Alcatel-Lucent

Key Words: Reverse regression, regression paradox

An analysis of customer survey data using direct and reverse regression led to some counter-intuitive conclusion. The phenomenon is known as the regression paradox. We describe sufficient conditions when the regression paradox will appear both analytically and geometrically. In linear models, it is a result of a distribution shift among groups relative to predictability in regression and reverse regression.

397 Inference and Estimation

Section on Statistical Computing

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

A Generalized Single-Stage Algorithm To Compute Shortest Confidence Intervals

✱Ming Yang, The University of Iowa, Biostatistics Department, 421 Hawkeye Court, Iowa City, IA 52246, *ming-yang@uiowa.edu*; Richard Dykstra, The University of Iowa

Key Words: confidence level, pivotal quantity, shortest confidence intervals

Shortest confidence intervals have not received a great deal of attention largely due to the difficulty in computation. A two-stage algorithm has been developed to solve the resulting equations derived from the constrained optimization. Unfortunately, the two-stage algorithm is a bit unwieldy and time-consuming. This paper introduces a generalized single-stage algorithm to obtain the solution in an efficient manner.

Statistical Power of Two Tests for Comparing Vaccination Coverage Curves, National Immunization Survey (NIS)

✱Zhen Zhao, Centers for Disease Control and Prevention, 1600 Clifton Road, N.E., MS E62, Atlanta, GA 30333, *zzhao@cdc.gov*

Key Words: Power, Curves, Test, Vaccination Coverage

We investigated the statistical power of two tests for comparing vaccination coverage curves. A Z-test was developed and applied to NIS data, but the power of this test has not been evaluated. In this paper we proposed a new Chi-square test, studied the test power for both the Z-test and the Chi-square test, and compared the power of two tests with different standard errors, mean differences, interaction measurements between curves, and controlling Type I error. If the two curves are approximately parallel, the power of Z-test is consistently higher than that of Chi-square test. For high standard error and low interaction, the power of Z-test is higher than that of Chi-square test. However, for low standard error or high standard error and high interaction, the power of Chi-square test is higher than that of Z-test. The curve comparison-wise power matrix among 20 areas was obtained also.

Assessment of Normality Tests for ANOVA and Multilevel Models Under Small Samples

✱Yi-Ting Hwang, National Taipei University, 67, Ming-Sheng E. Rd, Sec. 3, Taipei, 104 Taiwan, *hwangyt@mail.ntpu.edu.tw*; Chun-Chao Wang, National Taipei University

Key Words: ANOVA, multilevel model, normality test, W test, Neyman smooth test

Statistical models designed for continuous response variables are often constructed under the normal assumption. Bonett and Woodward (1990) indicated that the univariate normality test is sufficient to examine this assumption under large samples. In some studies, due to cost or time constraints, the sample size may be small. In addition, many survey data are collected under rather complex sampling designs. Some data require the use of multilevel models. Hwang and Wei (2007) showed that the univariate test cannot be used unless a transformation is performed. Simulations are conducted to evaluate the feasibility of normality tests under the ANOVA model and the two-stage model when small samples are considered.

Comparison of the Test for Normality Using Goodness-of-Fit Test Based on the Likelihood Ratio

✱Ampai Thongteeraparp, Kasetsart University, 50 Phaholyothin Rd., Chatuchak, Department of Statistics, 10900 Thailand, *fsciamu@ku.ac.th*; Sirithip Wasinrat, Chandrakasem Rajabhat University

Key Words: Goodness of fit test for normality, Likelihood ratio, Power of the test, Shapiro-Wilk Statistic, Shapiro-Francia Statistic

The purpose is to compare power of the test of Kolmogorov-Smirnov statistic based on the likelihood ratio(Zk), Anderson-Darling Statistic based on the likelihood ratio(Za), Cramer von Mises Statistic based on the likelihood ratio(Zc), Shapiro-Wilk(W1) and Shapiro-Francia(W2)in testing normality. The studied distribution are normal, t, chi-square, beta and log-normal distribution. The sample sizes are 10,30,50 and 100. The specified significance levels are .01,.05 and.10. Using Monte Carlo simulation technique, the studies show that Zk, Za and Zc can control the observed significance levels for every sample sizes. W1 cannot control the observed significance levels when sample size is 100. W2 cannot control the observed significance levels for every sample sizes. For power of the test, most cases of power of W1 and W2 are high at every distribution, sample sizes and significance levels.

Estimating Reliability and Comparing Its Estimators

✱Mohammed A. Shayib, Prairie View A&M University, Department of Mathematics, P. O. Box 519 MS 2225, Prairie View, TX 77446-0519, *mashayib@pvamu.edu*; Aliakbar M. Haghighi, Prairie View A&M University

Key Words: Reliability, Probability Distributions, Methods of Estimation, Populations' Parameters, Sample Sizes, Simulation

The estimation of reliability R = P(Y<X), assuming X and Y to be independent and identically distributed random variables, has been extensively studied in the literature. Distribution functions such as Normal, Burr Type X, Exponential, Gamma, and Weibull have been considered. In this paper we consider the Weibull type of distri- butions. Utilizing the Maximum LikelihoodEstimation (MLE),Uniformly Minimum Varuance Unbiased Estimation (UMVUE), Shrinkage Estimation Procedures, and the Method of Moments, using simulation, we will estimate the unknown parameters of the Weibull distributions. Among the displayed estimators for assumed distributions, found by the different methods, a comparison will be made to choose that estimator(or estimation method) which gives the best results in calculating R. Those results will be made based on the ratio of the populations' parameters and sample sizes.

Tolerance Limits for the Distribution of the Difference Between Two Independent Normal Random Variables

✱Sumona Mondal, Clarkson University, Department of Mathematics, P O Box 5851, Potsdam, NY 13699, *smondal@clarkson.edu*; Xiadong Lian, University of Louisiana at Lafayette; K. Krishnamoorthy, University of Louisiana at Lafayette

Key Words: Coverage probability, Lognormal distribution, Moment approximation, Stress-strength reliability

The problem of constructing tolerance interval for the distribution of the difference between two independent normal random variables possibly with different variances is considered. Assuming that the variance ratio is known, an exact method and a simple approximate method for constructing tolerance factor are presented. Also, an approximate method for obtaining tolerance interval is provided when the variances are unknown and arbitrary. The accuracy of the approximate method is evaluated via Monte Carlo simulation study. Simulation study indicates that the approximate tolerance interval is quite satisfactory when the sample sizes are small and close to each others or moderate and not drastically different. The proposed tolerance interval can be used safely for practical applications. The methods are illustrated using two examples.

398 Advances in Computer Experiments ●

Section on Physical and Engineering Sciences
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Two-Stage Group Screening for Computer Experiments

✱Hyejung Moon, The Ohio State University, Department of Statistics, 1958 Neil Avenue, Columbus, OH 43210-1247, *hjmoon@stat.osu.edu*; Thomas J. Santner, The Ohio State University; Angela M. Dean, The Ohio State University

Key Words: computer experiment, two-stage, group screening, sensitivity index, Latin hypercube design, high-dimensional

In computer experiments with high-dimensional inputs, a single run of the computer code can be extremely time consuming and rapid identification of important inputs is crucial. A two-stage group screening procedure is proposed which first partitions individual inputs into groups and fits a model to the groups. Active groups are identified and inputs in these groups proceed to the second stage. After collecting additional data, active individual inputs are determined. Decisions about active inputs are based on a yardstick constructed from a sensitivity index distribution of the dummy input. The best orthogonal array-based Latin hypercube design under a minimax correlation criterion is used in each stage. Simulation studies show the two-stage procedure can be better than one-stage procedure in terms of error rates, number of runs, and computation times.

Simultaneous Determination of Calibration and Tuning Parameters

✱Gang Han, The Ohio State University, 1958 Neil Avenue, Cockins Hall,, Columbus, OH 43210, *han.191@osu.edu*; Thomas J. Santner, The Ohio State University

Key Words: Gaussian stochastic process model, Product Exponential correlation functions, Metropolis Hastings algorithm, the Law of Large Numbers

Increasingly, scientists and engineers use computer experiments as surrogates or supplements to the physical experiments. Inputs to computer experiments include engineering design variables, calibration parameters (inputs having unknown true values and physical meanings), and tuning parameters (variables typically controlling code's solution tolerances and having no physical meanings). Calibration and tuning parameters are inputs only for computer experiments. Operationally, one could treat both calibration and tuning parameters as calibration parameters. This talk introduces a methodology for simultaneous calibration and tuning in settings where data are available from both a computer experiment and an associated physical experiment. Our method is based on a Bayesian model and a Monte Carlo Markov Chain simulation algorithm. We compare our methodology with an alternative approach.

Sequential Calibration of Computer Models

✱Arun Kumar, The Ohio State University, 304C, Cockins Hall, 1958 Neil Ave., Columbus, OH 43210, *arunk@stat.osu.edu*; William Notz, The Ohio State University

Key Words: Computer Experiments, Calibration, Sequential Designs, Gaussian Process Models, Space Filling Designs

We propose a sequential method for the estimation of calibration parameters for computer models. In this method, we first fit separate Gaussian Stochastic process (GASP) models to given data from a physical and a computer experiment. The value of the calibration parameters that minimizes the discrepancy between predictions from the two models, are taken as the estimates. In the second step, the point with maximum potential for reducing the uncertainty in the fitted model is identified. Computer experiment is conducted at this new point. The first step is repeated with the augmented data set, the calibration parameters re-estimated, and the next design point determined. The method is repeated until the allocated budget for the number of design points are exhausted. Empirical results are presented to show effectiveness of the method in achieving faster convergence to the estimates.

Modeling and Estimation of Physical Simulation Parameters in Data Assimilation

✱Kazuyuki Nakamura, The Institute of Statistical Mathematics, International Japan, *nakakazu@ism.ac.jp*; Tomoyuki Higuchi, The Institute of Statistical Mathematics

Key Words: data assimilation, particle filter, state space model

Data assimilation is developed in the meteorology and the oceanography, which combines states of a physical simulation model with observations. It aims at providing a good initial condition for the nonlinear simulation model and/or estimating model parameter online. In data assimilation, ensemble based filters such as the ensemble Kalman filter and the particle filter are employed. At first, we discuss assimilation performance among ensemble based filters through Lorenz 63 and 96 models. In next, we introduce estimation framework of unknown parameters in simulation model, which is regarded as estimation of hyper-parameters of the model. Through the numerical experiment using Lorenz model, we also discuss how to model these unknown parameters.

Dynamically Consistent Sequence-Based Data Assimilation

✱Neil Martinsen-Burrell, Wartburg College, MCSP Department, 100 Wartburg Blvd, Waverly, IA 50677, *nmb@wartburg.edu*

Key Words: data assimilation

Traditional approaches to data assimilation use sequential observations of a system to make a best guess about the current state of a dynamical system. In this talk, we use a sequence of observations to reconstruct a sequence of system states. We do this by minimizing a convex combination of the mismatch between the state and the observations and the mismatch between each state and our prediction for that state. Our experiments use a simple two vortex model with a single passive drifter. Using a simple gradient-descent process, we reconstruct sequences of system states that are consistent with the observational noise model used to generate the observations. We also reconstruct consistent sequences of system states using observations only of the position of the drifter.

399 Bayesian Computational Statistics ●

Section on Bayesian Statistical Science, Section on Statistical Computing

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

New Perspectives on the Estimation of Normalizing Constants via Posterior Simulations

✱Giovanni Petris, University of Arkansas, SCEN 301 Mathematical Sciences, 1 University of Arkansas, Fayetteville, AR 72701, *GPetris@uark.edu*; Luca Tardella, Universit‡ di Roma "La Sapienza"

Key Words: Normalizing constant, Integrated likelihood, MCMC

In this paper we recast the problem of evaluating the normalizing constant of an arbitrary density function with the aid of an arbitrary MC or MCMC sampling scheme. One new basic idea is that of perturbing the original target density function, whose normalizing constant has to be evaluated, in such a way that the perturbed density has the same original normalizing constant plus a known arbitrary positive mass. We focus on the effectiveness of the estimators derived from this new idea when the perturbed density is obtained via the Hyperplane Inflation idea of Petris and Tardella. The proposed estimators share the original simplicity of the harmonic mean estimator, yielding consistent MC or MCMC estimators based only on a simulated sample from the distribution proportional to the original density. The calibration of the optimal choice of the arbitrary positive mass is also discussed.

Geometric Ergodicity of Hybrid Samplers for Ill-Posed Inverse Problems

✱Radu Herbei, The Ohio State University, 1958 Neil Ave, Cockins Hall, Room 223, Columbus, OH 43210, *herbei@stat.osu.edu*; Ian W. McKeague, Columbia University

Key Words: hybrid samplers, geometric ergodicity, MCMC, PDE

Ill-posed inverse problems arise in many scientific fields, and statistical methods to address such problems have been studied extensively. In the Bayesian approach to inverse problems, regularization is imposed by specifying a prior distribution on the parameters of interest and MCMC samplers are used to extract information about its posterior distribution. The aim of the present paper is to investigate the convergence properties of the random-scan random walk Metropolis (RSM) algorithm, which is known to converge geometrically under certain conditions. We provide an accessible set of sufficient conditions, in terms of the observational model and the prior, to ensure geometric ergodicity of RSM samplers of the posterior distribution. We illustrate how these conditions can be easily checked in the context of an application to the inversion of oceanographic tracer data.

A Copula-Based Adaptive MCMC Sampler

✱Georges Tsafack, Suffolk University, 499 VFW Parkway, Chestnut Hill, MA 02467, *gtsafack@suffolk.edu*; Yves F. Atchade, The University of Michigan

Key Words: MCMC, Simulation, Copula

In Markov Chain Monte Carlo simulation, proper scaling of the so-called proposal kernel is important for good convergence properties. There is a recent trend in the literature to design adaptive MCMC algorithms where these proposal kernels are tuned automatically to their optimal values as the simulations unfold. These algorithms hold the potential of freeing the user from the tedious parameter tuning process of MCMC simulation. But currently, these samplers are limited in that they can only capture the linear dependence between the components of the target distribution. In this paper, working in a two-dimensional space, we show that much more versatility can be achieved using proposal kernel whose dependence structure is modeled by a copula. We propose an adaptive MCMC sampler that selects the best copula to fit the target distribution.

Partially Bayesian Variable Selection in Classification Trees

✱Douglas A. Noe, Miami University, Department of Mathematics and Statistics, 301 S Patterson Ave, Oxford, OH 45056, *noeda@muohio.edu*; Xuming He, University of Illinois at Urbana-Champaign

Key Words: feature selection, expert opinion, supervised learning

Currently available statistical models for classification may be split into two broad categories: (1) data-driven and (2) theory-based. Tree-structured classifiers such as CART and QUEST are members of the first category. In those algorithms, all predictor variables compete equally for a particular classification task. However, in many cases a subject-area expert is likely to have some qualitative notion about their relative importance. We introduce a partially Bayesian procedure for dynamically incorporating such qualitative expert opinions in the construction of classification trees. Such an algorithm has two potential advantages. First, unnecessary computational activity can be avoided. Second, we reduce the chance that a spurious variable will appear early in the model. Our models are potentially more interpretable and less unstable than those from purely data-driven algorithms.

Mechanism-Based Emulation of Dynamic Simulators: Concept and Application in Hydrology

✱Peter Reichert, Eawag, Duebendorf, International 8600 Switzerland, *reichert@eawag.ch*

Key Words: emulator, statistical inference, computationally demanding model

Emulators, by providing computationally efficient interpolation between outputs of deterministic simulation models, can considerably extend the range of statistical techniques that can be applied to computationally demanding models. So far, the dominant technique for developing such emulators were based on Gaussian stochastic process priors. In the context of dynamic models, this approach has two essential disadvantages: (i) these emulators do not consider our knowledge of the structure of the model, and (ii) they can run into numerical difficulties if the output covers the time dimension densely. To address these problems we propose an emulator that is based on a simplified linear state-space model with Gaussian stochastic processes of the innovation terms as functions of model parameters and/or inputs. Conditioning this prior to the design data set is done by Kalman smoothing.

Bayesian Boundary Restoration Using Splines

✱Larissa Stanberry, University of Washington, Department of Statistics, Box 354322, Seattle, WA 98195, *lstan@u.washington.edu*

Key Words: boundary detection, bayesian model, splines, mammogram

In image analysis, it is often desirable to reconstruct the boundary of an object in a noisy image. For example in mammography, boundary outline of the breast tissue is required as a part of a routine image assessment procedure to determine the risk of developing breast cancer. Various methods for boundary detection have been proposed (Markov random field models, active contour models, and stochastic models with polygonal shapes). This work introduces a novel method based on B-spline curves. The method allows to estimate a variety of complex contours, including boundaries of nonconvex sets and nonsmooth domains. The method is Bayesian and uses Markov chain Monte Carlo simulations to draw inference about the boundary. The method is illustrated on the simulated data and applied to recover a skin line on a digitized analog mammogram image. We also discuss the selection of a loss function.

Bayesian Transductive and Semi-Supervised Learning

✱Sounak Chakraborty, University of Missouri-Columbia, 209F middlebush hall, columbia, MO 65211, *chakrabortys@missouri.edu*

Key Words: transductive learning, prediction, bayesian analysis, diffuse prior, support vector machine, semi-supervised learning

This paper introduces a Bayesian semi-supervised support vector machine model and a transductive Bayesian nearest neighbor model for prediction and classification. The transductive or semi-supervised learning has a distinct advantage over the inductive learning since by formulation it removes the problem of over fitting. While a traditional SVM has the widest margin based on the labeled data only, our semi-supervised form of SVM attempts to find the widest margin in both the labeled and unlabeled data space. This formulation of semi-supervised learning enables us to use some information from the unlabeled data and improve the prediction performance. A penalty term is added for the parameters of the model and for also the likelihood constructed from the unlabeled data. The parameters and penalties are controlled thorough near diffuse priors for objectivity of the analysis.

400 Statistical Analysis of Time-Indexed Data

IMS, Section on Nonparametric Statistics
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

On Optimal Maximum Likelihood Estimation for Locally Stationary Long-Memory Processes

✱Jan Beran, University of Konstanz, Department of Mathematics and Statistics, University of Konstanz, Konstanz, International 78457 Germany, *jan.beran@uni-konstanz.de*

Key Words: long memory, fractional ARIMA process, local stationarity, bandwidth selection, maximum likelihood estimation

Parameter estimation for time-dependent locally stationary long-memory processes is considered. A limit theorem for a local maximum likelihood estimator is derived. Asymptotic formulas for the mean squared error lead to an asymptotic formula for the optimal bandwidth. Quite surprisingly, local estimation of \$d\$ turns out to be comparable to regression smoothing with iid residuals in the sense that the optimal bandwidth is of the order \$n^{-1/5}\$ and inversely proportional to the square of the second derivative of \$d\$. Several data examples illustrate the method.

On the Maximum of Some Processes with Negative Drift and Heavy Tail Innovations

✱William P. McCormick, The University of Georgia, Statistics Department, Athens, GA 30602, *bill@stat.uga.edu*

Key Words: maximum of random walk, heavy tail, fractional ARIMA, long range dependence

A classical result due to Veraverbeke gives the asymptotic tail behavior of the supremum of a random walk with heavy-tail innovations and negative mean. In this talk we will present extensions to this result where the innovations defining the random walk may be dependent. In particular we will allow for innovations constituting a nonstationary FARIMA process. More generally, we discuss how the single large jump heuristic applies in this setting. We also discuss sample path behavior conditionally on the supremum exceeding a large value.

Regularized Autoregressive Frequency Estimation

✱Bei Chen, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1 Canada, *beiaries@gmail.com*; Yulia R. Gel, University of Waterloo

Key Words: time series analysis, regularization, frequency estimation, autoregressive approximation, generalized spectral density, cross validation

Many real life processes can be modeled as a noised sum of sinusoids. Unknown frequencies can be obtained by approximating generalized spectral density of a process by an autoregressive (AR) model. The advantage is that an AR model has a simple structure, and its estimation and asymptotic properties are well-known. Usually, the order of the AR approximation is chosen by information criteria. However, with an increase of a sample size in on-line applications, model order may change, which leads to re-estimation of all parameters. We propose a regularized AR (RAR) approximation, which enables to estimate parameters with different level of accuracy. Hence, the repeated model selection and parameter estimation are avoided. Regularizer is chosen by cross-validation. In this talk, we discuss asymptotic properties of RAR, links to other regularization techniques and present numerical studies.

Asymptotic Properties of Hill’s Estimator for Shot Noise Sequence

✱Chenhua Zhang, Colorado State University, Department of Statistics, Fort Collins, CO 80523, *zhang@stat.colostate.edu*

Key Words: shot noise, heavy tail, regular variation, tail index, Hill’s estimator

We study a shot noise sequence of the form $X_j=\sum_{i \leq j} h(\tau_j-\tau_i)A_i$, where the distribution of A_i has regularly varying tail with index $-\alpha$. Under a mild integration condition on h the distribution of X_i also has regularly varying tail with index $-\alpha$ and the Hill’s estimator is a weakly consistent estimator of $1/\alpha$. If the tail distribution of X_i is normalized and second order regularly varying the Hill’s estimator is asymptotically normal for a large class of functions h.

Random Partition Masking Model for Censored and Masked Competing Risks Data

✱Qiqing Yu, Binghamton University, Math Department, Binghamton, NY 13902, *qyu@math.binghamton.edu*; Hao Qin, Binghamton University; Jiaping Wang, Binghamton University

Key Words: Right-censoring, competing risks model, masking, symmetry assumption, MLE, consistency

We consider the parametric estimation with right-censored competing risks data and with masked failure cause. We propose a new model, called the random partition masking model. We propose a wide class of parametric distribution families of the failure time and cause. A feasible algorithm for the MLE is proposed for some special cases. We also study the asymptotic properties of the MLE under the new model.

401 Robust Estimation for Skewed and Heavy-Tailed Distributions

Biometrics Section, Section on Nonparametric Statistics

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

The Epsilon-Skew Laplace Distribution

✱Hassan Elsalloukh, University of Arkansas at Little Rock, Department of Mathematics and Statistics, 2801 South University Avenue, Little Rock, AR 72204-1099, *hxelsalloukh@ualr.edu*

Key Words: Skewness, Kurtosis, Skew-Laplace, Skew-Normal

In this paper, I introduce a new distribution family that I name the Epsilon-Skew Laplace Distribution (ESL). This set of random variables possesses both symmetric and asymmetric density functions and includes the Laplace distributions as a special case. I define basic properties and highlight special members of the ESL distribution family. I also derive general expressions for the mean, variance, skewness, kurtosis, moments about zero and about a location parameter, maximum entropy, univariate generator, Fisher Information, score tests, and maximum likelihood estimation for the three parameters involved in the new distribution.

Three-Step Estimation in Linear Mixed Models with Skew-T Distributions

✱Tianyue Zhou, sanofi-aventis, Hillsborough, NJ , *tianyue.zhou@sanofi-aventis.com*; Xuming He, University of Illinois at Urbana-Champaign

Key Words: Linear mixed models, Random effects, Skew-t

Linear mixed models based on the normality assumption are widely used in health related studies. Although the normality assumption leads to simple, mathematically tractable, and powerful tests, violation of the assumption may easily invalidate the statistical inference. Transformation of variables is sometimes used to make normality approximately true. In this paper we consider another approach by replacing the normal distributions in linear mixed models by skew-t distributions, which account for skewness and heavy tails for both the random effects and the errors. The full likelihood-based estimator is often difficult to use, but a 3-step estimation procedure is proposed, followed by an application to the analysis of deglutition apnea duration in normal swallows. The example shows that skew-t models often entail more reliable inference than Gaussian models for the skewed data.

On the Use of Optimal Discovery Procedure in Skewed Distributions

✱Igor Melnykov, Colorado State University, Pueblo, Department of Mathematics and Physics, 2200 Bonforte Blvd., Pueblo, CO 81001, *igor.melnykov@colostate-pueblo.edu*

Key Words: multiple testing, optimal discovery procedure, unbiased tests

The paper investigates the properties of the optimal discovery procedure (ODP) applied in multiple hypothesis testing. In particular, we consider a testing situation where test statistics have positively skewed distributions. The performance of the ODP is compared to that of traditional testing techniques. We also present some results on the unbiasedness of the ODP.

Heavy-Tailed Distributions in Quantitative Trait Genetics

✱Serge Sverdlov, University of Washington, Box 354322, Department of Statistics, Seattle, WA 98195-4322, *serges@stat.washington.edu*

Key Words: QTL, polygenic traits, heavy tailed distribution, cauchy distribution, population genetics, statistical genetics

The statistical genetics of quantitative traits have benefited deeply from analysis of variance methods, which carry assumptions of finite moments and normal kurtosis. Heavy tailed distribution methods violate these assumptions and can therefore lead to radically different behaviors. Our assertion is that the genetics of heavy-tailed traits is not merely Gaussian genetics with outliers; rather, these traits exhibit subtle heritability patterns, especially in polygenic traits, and show the classical distinction between heterogeneous Mendelian and blending forms of inheritance by a novel bifurcation argument. We derive closed form results using the Cauchy distribution, which we show to be a steady state in a broad family of population processes, and connect the distribution to evolvability via extreme value theory and process subordination arguments.

Robust Estimation for Finite Mixture Regression Models with Random Effects via the Minimum Hellinger Distance Approach

✱Liming Xiang, Nanyang Technological University, Singapore, 101C Nanyang View, #11-05, Singapore, 639670 Singapore, *lmxiang@ntu.edu.sg*

Key Words: Kernel estimator, Minimum Hellinger distance, Finite mixture, Residual maximum likelihood estimation, Robustness

The finite mixture regression with random effects is effective in modeling simultaneously the heterogeneity for clustered data arising from several latent subpopulations and the dependency among the data. In the literature, the residual maximum likelihood estimation (REML) has been studied for the model. However, the REML estimators tend to be unstable in the presence of outliers or extreme contamination. This paper focuses on developing robust estimators using the minimum Hellinger distance (MHD) approach. Under certain conditions, the MHD estimators are shown to be consistent and asymptotically normal. Monte Carlo simulations show that the MHD estimators perform satisfactorily for data without outlying observation(s), and outperform the REML estimators when data are contaminated. Application to a data set of neonatal length of stay (LOS) is presented to illustrate the method.

A New Method for Estimating Regression Parameters with Repeated Runs

✱Howraa Majeed, University of Arkansas at Little Rock, Department of Applied Science, ETAS 575, 2801 South University Avenue, Little Rock, AR 72204, *hamajeed@ualr.edu*; Hassan Elsalloukh, University of Arkansas at Little Rock; Shawki Shaker Hussain, University of Baghdad

Key Words: Repeated Runs, Two Phases-Method, M - Estimates

In this paper, we deal with the problem of estimating the slop and intercept parameters of the simple linear regression model in the presence of outliers when repeated runs of observations are available. For such a problem, we propose various robust estimators that are based on a new method called, "Two Phases-Method." In Phase-I, we simply employ, at each fix value of the independent variable, one of the robust methods of estimation of the location parameter to estimate the unknown value of the dependent variable. At Phase-II, the reduced new observations obtained by Phase-I will be used to estimate the slope and intercept parameters by either the ordinary least squares method or, again, by one of the robust methods. Simulation results show the superiority of the suggested method of the usual least squares or robust method.

Estimation of False Discovery Rate Using Skew-Mixture Models

✱Anindya Roy, University of Maryland, Baltimore County, 100 Hilltop Circle, Baltimore, MD 21250, *anindya@math.umbc.edu*; Subhashis Ghosal, North Carolina State University

Key Words: Multiple Testing, Mixture Model

We propose a skew-mixture model for the p-value density and use the model to estimate important functionals such as the positive false discovery rate. The model admits shape restrictions in the p-value density commonly observed in practice.

402 Models, Measurements, and Methods

General Methodology

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Optimizing Smoothness Kernel Size for the Smoothed Variance T-Test

✱Hui Zhang, The University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109, *huizhang@umich.edu*; Timothy D. Johnson, The University of Michigan; Thomas E. Nichols, GlaxoSmithKline Clinical Imaging Centre

Key Words: smoothness parameter, smoothed variance image, Bias, MSE, effective degrees of freedom, Chisquare random field

For small group studies with fewer subjects, the smoothed variance t-test has been found to be a powerful alternative to the usual t-test. The power is greater because smoothing the variance effectively increase the error degrees of freedom (DF). However, the unsmoothed sample variance image is unbiased and so any smoothing will induce some bias, the severity of which depends on the smoothness of the true variance image. In this work we develop a theoretical framework to obtain bias, mean square error (MSE) and effective DF (EDF) of the smoothed variance image, as a function of data smoothness, true variance image smoothness and applied variance smoothing. One novel aspect is our use of Chi-square random fields to model the true variance. We use simulation to evaluate this method. We also use splits a 150 subject data set to find equivalent empirical results.

Normal-Based Methods for a Gamma Distribution: Prediction and Tolerance Intervals and Stress-Strength Reliability

✱Shubhabrata Mukherjee, Virginia State University, Department of Mathematics and Computer Science, 1 Hayden Drive,, Petersburg, VA 23803, *smukherjee@vsu.edu*; K. Krishnamoorthy, University of Louisiana at Lafayette; Thomas Mathew, University of Maryland Baltimore County

Key Words: Tolerance Limits, Coverage probability, Survival probability, Wilson-Hilferty approximation

In this article, we propose inferential procedures for a gamma distribution using the Wilson-Hilferty normal approximation. We propose normal based approaches for a gamma distribution (i) for constructing prediction limits, one-sided tolerance limits and tolerance intervals, (ii) for obtaining upper prediction limits for at least l of m observations from a gamma distribution at each of r locations, and (iii) for assessing the reliability of a stress-strength model involving two independent gamma random variables. For each problem, a normal based approximate procedure is outlined, and its applicability and validity for a gamma distribution are studied using Monte Carlo simulation. Our investigation shows that the approximate procedures are very satisfactory for all of the above problems. For each of the problems considered, the results are illustrated using practical examples.

The Effect of Toeplitz Structure on the Power of Multivariate Significance Tests in a MANOVA

✱Jessica H. Scott, Brigham Young University, 1001 SWKT, Psychology Department, Provo, UT 84602, *jessiemae2000@aol.com*; Bruce L. Brown, Brigham Young University

Key Words: MANOVA, statistical power, Toeplitz structure, clustered structure

In conducting a multivariate analysis of variance (MANOVA), Rencher & Scott, 1990, recommend using the four multivariate tests of significance to screen the corresponding univariate tests, as an effective method for controlling alpha inflation. One would perhaps expect that the vector configuration of the dependent variables in such an analysis would have no effect upon the power of the multivariate significance tests, but such is not the case. Using Monte Carlo simulation methods a variety of between/within multivariate data configurations are created, half of them with Toeplitz vector patterns for the dependent variables, and half with clustered vector patterns. The multivariate significance tests are found to be substantially less powerful when the vector patterns are Toeplitz than when they are clustered.

Exact Bounded Risk Estimation When the Terminal Size and Estimator Are Dependent: The Gamma Case

✱Kevin P. Tolliver, Auburn University, Department of Mathematics and Statistics, Auburn, AL 36830, *kptolliver@gmail.com*; David M. Carpenter, Auburn University

Key Words: sequential, bounded risk estimation, optimal sample size, gamma distribution

It has been previously shown a purely sequential method for approximating the optimal sample size is better than other sequential methods sequential, for example, the exponential case. In this paper, we develop a purely sequential risk estimation procedure for estimating the mean of a gamma process. Because, in this case, the sequential procedure uses the combined sample from all stages, the estimator and the sequential sample size are dependent. We show that our method guarantees that we have an exact bounded risk estimate. These methods are demonstrated through extensive simulation studies and application to real data sets.

An Empirical Approach to Sufficient Similarity in Dose-Responsiveness: Utilization of Statistical Distance as a Similarity Measure

✱Scott Marshall, Virginia Commonwealth University, 730 East Broad Street, Department of Biostatistics, 3039b, Richmond, VA 23298, *marshallsl2@vcu.edu*; Chris Gennings, Virginia Commonwealth University; LeAnna G. Stork, Monsanto Company; Linda Teuschler, NCEA-U.S. EPA; John Libscomb, NCEA-U.S. EPA; Mike DeVito, U.S. Environmental Protection Agency; Kevin M. Crofton, U.S. Environmental Protection Agency

Key Words: risk assessment, chemical mixtures, non-linear mixed models, equivalence testing

Using statistical equivalence testing logic and mixed model theory an approach has been developed, that extends the work of Stork et al (*JABES*,2008), to define sufficient similarity in dose-response for chemical mixtures containing the same chemicals with different ratios or a subset of chemicals. The similarity measure is based on statistical distance which is defined as a function of relative potencies of the chemicals. A simulation study was conducted to assess the "power" of the approach. The mixing ratio for exposure to a chemical mixture was assumed to be a random process. The current work estimated how often a resulting candidate mixture was sufficiently similar in dose-responsiveness to the reference mixture. (Partially supported by NIEHS #T32 ES007334 and does not reflect USEPA policy. This research is not associated with Monsanto.)

403 Multiple Comparisons and Simultaneous Inference Methods

Biometrics Section, Biopharmaceutical Section, IMS
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

A Geometric Interpretation of Permutation P-Value and Its Application

✱Wei Sun, The University of North Carolina at Chapel Hill, 3104E McGavran-Greenberg Hall, Chapel Hil, NC , *wsun@bios.unc.edu*; Fred Wright, The University of North Carolina at Chapel Hill

Key Words: permutation, permutation p-value, linkage, eQTL, genetical genomics

Permutation p-value has been widely used to access the significance of linkage score in genetic studies. However its application in large-scale linkage studies is hindered by its heavy computational burden. In this study, we propose a method to estimate permutation p-values for linkage studies based on a geometric interpretation of permutation p-values. Suppose a quantitative/qualitative trait and p genetic markers are studied in n individuals. The p marker genotype profiles are p points in the n dimensional space. Permutation p-value is related with the space occupied by the p points. We define a new concept, "effective volume," which is the volume occupied by the observed genotype profiles according to the cutoff of linkage score, and develop an efficient algorithm to estimate the "effective volume," hence the permutation p-value.

Estimating the Proportion of True Null Hypotheses Using the Pattern of the Observed P-Values

✱Tiejun Tong, University of Colorado, Boulder, Department of Applied Mathematics, 526 UCB, Boulder, CO 80309, *tiejun.tong@colorado.edu*; Zeny Feng, University of Guelph; Hongyu Zhao, Yale University School of Medicine

Key Words: gene expression data, multiple testing, proportion of proportion of true null hypotheses, p-value

Estimating the proportion of true null hypotheses, π_0, has attracted much attention in the recent statistical literature. Most existing methods for estimating π_0 in the literature are motivated from the independence assumption of test statistics, which is often not the truth in reality and the performance in the presence of the dependence among test statistics can be poor. In this article, we propose a data-driven method for estimating π_0 by incorporating the distribution pattern of the observed p-values as a practical approach to addressing potential dependence among test statistics. The proposed estimators may substantially decrease the variance and thus improve the overall performance.

Monotone Multiple Testing Procedures and Implications of Monotonicity

✱Alexander Y. Gordon, The University of North Carolina, Charlotte, 9201 University City Blvd, Charlotte, NC 28223, *aygordon@uncc.edu*

Key Words: multiple testing procedure, monotone multiple testing procedure, family-wise erorr rate, generalized family-wise erorr rate

We consider the class of p-value-only based (or uninformed) multiple testing procedures (MTPs). We assume that an MTP is symmetric and cutting (significant p-values are smaller than insignificant ones). Monotonicity means that by replacing some p-values with smaller ones we can only increase the number of rejected null hypotheses. Several other, seemingly different, definitions of monotonicity turn out to be equivalent to this one. The property of monotonicity is natural and quite general; the vast majority of MTPs presented in the literature have it. At the same time, this property is quite strong; it has strong implications for the levels at which the MTP controls the family-wise error rate (FWER) and the generalized FWERs of different orders. Those implications will be discussed in the talk.

Simultaneous Confidence Bands for Multidimensional Effective Doses in Chemical Mixture Models

✱Gemechis D. Djira, South Dakota State University, Harding Hall 123, Brookings, SD 57007, *gemechis.djira@sdstate.edu*; Din Chen, South Dakota State University

Key Words: Chemical mixture, Effective doses, Joint action, Simultaneous inference

In dose-response experiments involving two or more chemicals, it is often of interest to study the joint action of the chemicals. The case of experiments with binary response variable and one chemical (drug) is studied extensively with emphasis on ED50 estimations. If there are more than one chemical involved, the joint action of these chemicals are of particular interest to test for synergism, antagonism or additivity, where the effective doses will be multidimensional. In this talk, we develop techniques for constructing simultaneous confidence bands for the effective doses as an alternative to point-wise confidence intervals based on the delta and Fieller intervals. The simultaneous coverage probabilities of the proposed methods will be compared with that of the existing procedure based on the inversion of Scheffe-type simultaneous confidence intervals.

Adaptive Choice of the Number of Bootstrap Samples in Large-Scale Multiple Testing

✱Wenge Guo, National Institute of Environmental Health Science, Durham, NC 27709-2233, *guow@niehs.nih.gov*; Shyamal Peddada, National Institute of Environmental Health Science

Key Words: bootstrap, confidence interval, false discovery rate, multiple testing, Benjamini-Hochberg procedure

It is a common practice to use resampling methods such as bootstrap for calculating p-value for each test when performing large scale multiple testing. However, the required number of bootstraps can be computationally burdensome. Further adding to the computational challenge is that the calculation of the test statistic itself may require considerable computation time. As technology improves one can expect the dimension of the problem to increase as well. Motivated by this important need, we developed a simple adaptive bootstrap methodology for large scale multiple testing, which reduces the total number of bootstrap calculations while ensuring the control of the FDR. Based on simulation study we found that, relative to the number of bootstraps required for the Benjamini-Hochberg procedure, the proposed methodology achieved a very substantial reduction in the number of bootstraps.

A Higher Dimension Multiplicity Adjustment Method When Testing for Multiple Response Variables

✱Mohammad Quasem, Howard University, Department of IS&DS, School of Business, 2600 6th Street, NW, Washington, DC 20059, *mquasem@howard.edu*; Mohammad F. Huque, U.S. Food and Drug Administration; Atiar Rahman, U.S. Food and Drug Administration

Key Words: Multiplicity Adjustments, inflation, Type I error rate, correlated response variables, ad-hoc methods, modification

Scientific experiments perform multiple statistical tests for various reasons, mostly to determine if an experimental treatment is better than a control. These multiple statistical tests, however can easily inflate the probability of false positive findings. This error probability is known as the family-wise Type I error rate and is often referred to as FWER. A few convenient methods for controlling this error rate have been studied. These studies are attractive but can be quite conservative in declaring statistical significance, particularly in situations when the test statistics for multiple response variables are highly correlated. Huque, Quasem, and Rahman improved the prior multiplicity adjustment method; they have found an improved FWER control. In this study, Huque, Quasem, and Rahman's adjustment method will be extended for a higher dimension of correlated response variables.

Surrogate Variable Analysis

✱Jeffrey Leek, Mount Sinai School of Medicine, One Gustave L. Levy Place, Box 1496, New York, NY 10029, *jeffrey.leek@mssm.edu*; John Storey, Princeton University

Key Words: multiple testing dependence, surrogate variables, gene expression, principal component analysis

High-throughput experiments measure data for thousands of related features and seek to rank them for association with variables of experimental or clinical importance. The process of ranking features for association with primary variables is complicated by other factors that may influence thousands of features at a time. In high-dimensional experiments these factors are often unknown, unmeasured, or incapable of being tractably modeled. We provide a statistical framework for modeling large-scale noise dependence caused by unmodeled factors in high-throughput data. We introduce surrogate variables, estimable linear combinations of the true unmodeled factors, that can be included when modeling the relationship between the primary variables and the feature data. We propose algorithms for estimating surrogate variables based on principal component analysis of subsets of features.

404 Sample Size Issues in Drug Research and Development ●▲

Biopharmaceutical Section, Biometrics Section
Wednesday, August 6, 10:30 a.m.–12:20 p.m.

Bayesian Sample Size Determination for Confirmatory Trial

✱Qin Pan, ICON Clinical Research, 555 Twin Dolphin Drive, Redwood city, CA 94065, *panq@iconus.com*; Charles DuMond, ICON Clinical Research

Key Words: confirmatory trial, power prior, sample size determination

Oncology trial often need take long time to get the result due to the competition, small patient population and long follow up time when the primary end point is time to event. FDA usually required confirmatory trial to approve a drug. We explore the application of design the second trial using Bayesian power prior method that presented by Santis(2007) to do sample size determination based on first trial result. Comparison will be made to the sample size for the second trial by using the full weight given to the prior trial result, discount weight to prior trial result, and classical frequency approach. The hypothetic confirmatory study design will use simulated data.

A Multivariate Approach to Sample Size Calculations for Thorough QT Studies

✱Sharon C. Murray, GlaxoSmithKline, 114 Springdale Way, Chapel Hill, NC 27517, *sharon.c.murray@gsk.com*; Suraj P. Anand, North Carolina State University; Gary G. Koch, The University of North Carolina at Chapel Hill

Key Words: QTc, Monte Carlo simulation, Thorough QT Study, correlated measurements

Thorough QT (TQT) studies are routinely conducted by pharmaceutical companies to satisfy the requirements of the ICH E14 Guidance. The cost involved in running a TQT study is substantial and sample size calculations play an important role in ensuring a small but adequately powered TQT study. In a TQT study, multiple observations are collected over time and it is reasonable to assume that observations are correlated and follow the multivariate normal distribution. We use Monte Carlo simulation to determine sample size varying the true maximum difference, the pattern of true differences over time, and the correlation between observations assuming a compound symmetric model for both the parallel and crossover designs. For the crossover design we also consider a variance components model in which correlation is greater for observations taken during the same period.

Sample Size for Identifying Differentially Expressed Genes in Microarray Experiment

✱Zhaoyu Luo, Rutgers, The State University of New Jersey, 149 Nichols Apts., Piscataway, NJ 08854, *zyluo@stat.rutgers.edu*; Javier Cabrera, Rutgers, The State University of New Jersey; Xiang Yu, Merck Reserach Laboratories; Peggy Wong, Merck Research Laboratories

Key Words: Microarray, sample size, clustering, differential expression, power, covanrance correction

Sample size calculation in the identification of differentially expressed genes is an important problem of microarray experiment. In this paper, we induce clustering and correlation correction to gene expression identification. On one hand, after a cluster analysis, genes performing similar functions or participating in the same genetic pathway would congregate in the same cluster. On the other hand, various researchers have suggested that accounting for correlation among genes could improve the power. There exists, however, three challenges: the first one is which clustering method we choose, the second one is how to reliably estimate the covariance matrix of genes and the last one is how to model the data, perform the appropriate statistical test and calculate the power. Here we present our methodology to tackle the problems.

The Minimum Required Number of Clusters for Cluster Randomized Trials

✱Zhiying You, The University of Alabama at Birmingham, *zyyou@uab.edu*; Gary Cutter, The University of Alabama at Birmingham

Key Words: cluster randomized trial, sample size, statistical power, feasibility

There are two sample size considerations in cluster randomized trials (CRTs): the number of clusters per intervention group and the number of individuals in each cluster (called cluster size). We show the unequal contribution of the number of clusters and cluster size to study power. The former is the main determinant though the later plays an important role. We also show the existence of an upper bound of study power for a given number of clusters per group when the cluster size tends to infinity, and thus the existence of a minimum number of clusters per group for a prespecified study power. The minimum required number of clusters per group actually serves as a criterion for the feasibility of CRTs. This is very important especially for trials with small number of clusters of large cluster size. We develop methods to calculate the sample size requirements and the upper bound of power.

Sample Size Calculation for an Agreement Study

✱Jason Liao, Merck Research Laboratories, P.O.Box 4, West Point, PA 19486, *jason_liao@merck.com*

Key Words: Concordance, Agreement interval, Sample size, Discordance rate, Tolerance probability

It is often necessary to compare two measurement methods in medicine and other experimental sciences. This problem covers a broad range of data. Many authors have explored ways of assessing the agreement of two sets of measurements. However, there has been relatively few attention to the problem of determining sample size for designing an agreement study. In this paper, a method using the interval approach for concordance is proposed to calculate sample size in conducting an agreement study. The sample size calculation is based on two rates: the discordance rate and tolerance probability, which in turn can be used to quantify an agreement study. The proposed approach is demonstrated through a real data set.

An Optimal Group Size Ratio for Many-to-One Comparisons in Clinical Trials

✱Jianliang Zhang, MedImmune, Inc., One Medimmune Way, Gaithersburg, MD 20878, *zhangj@medimmune.com*

Key Words: optimal design, many-to-one comparison

It is common in a proof of concept trial to compare multiple active treatments with a common control. The active groups are usually of equal size and the control group is a ratio of the active group size. This study finds the theoretical optimal active-to-control ratio with respect to the efficiency in the use of statistical information and power of individual active to control comparisons. The efficiency of a design with an active-to-control ratio is compared to an equally-sized design. For a study endpoint that follows a normal distribution with equal variance, the optimal ratio is a function of the number of groups (K) in the study. For a binary endpoint, the optimal ratio becomes a function of K and a binary proportion (p). An average efficiency over an interval of p is used to determine the optimal active-to-control ratio in the design of such studies.

On Adaptive Trial Simulations

✱William W. Wang, Merck & Co., Inc, 381 Huai Hai Middle Road, Shanghai Central Plaza, 26 floor, Shanghai, 200020 P.R. China, *william_wang@merck.com*

Key Words: Adaptive Design, Clinical Trial Simulation

Adaptive trial design provides a promising approach to simultaneously improve decision making and saving time and money in drug development, without compromising quality. Such design, however, results in a much more complex matrix of planning design options. In order to successfully validate and optimize the considered options, performing simulations that reflect the complex situations is critical and challenging. In this talk we argue that, similar to a real clinical study, a simulated adaptive study should be carefully planned, executed, analyzed and reported. This talk further details some important considerations. These include (1) setting simulation objective; (2) designing simulation frameworks & process flows; (3) deciding proper metrics to assessing operating characteristics; (4) choosing realistic yet tangible simulation scenarios; and (5) implementing in a robust way.

405 Contributed Poster Presentations

Biometrics Section, Biopharmaceutical Section, Business and Economics Statistics Section, ENAR, IMS, Section on Nonparametric Statistics, Section on Quality and Productivity, Section on Statistical Computing, Section on Statistical Graphics, Section on Statisticians in Defense and National Security, Section on Statistics and Marketing, Section on Statistics and the Environment, Section on Statistics in Epidemiology, WNAR

Wednesday, August 6, 10:30 a.m.–12:20 p.m.

An Application of Multiple Partial Imputation to Analysis of Longitudinal Quality-of-Life Data

✱Paul Kolm, Christiana Care Health System, 131 Continental Drive, Suite 202, Newark, DE 19713, *pkolm@christianacare.org*; Wei Zhang, Christiana Care Health System; John Spertus, Mid American Heart Institute; William Weintraub, Christiana Care Health System

Key Words: Missing data, Longitudinal data analysis, Multiple partial imputation

Missing data present a challenge for analysis, particularly when the data are missing not at random (MNAR). In clinical trials where longitudinal quality of life and health status data are obtained from patient self-report, missing observations may be MNAR, in that patients in poorer health may be reluctant to complete a questionnaire regarding their health. A multiple partial imputation (MPI) strategy has been suggested that assumes intermittent missing data are missing at random (MAR), but missing data due to dropout are MNAR (nonignorable). The imputed data sets are then analyzed by likelihood or quasi-likelihood methods that account for nonignorable dropout. In this study, MPI is applied to quality of life data obtained in a large clinical trial spanning a 7 year follow-up period. Issues relating to the assumptions and application of MPI are discussed.

Imputation of Missing Items for the Scale Variables Using Item Response Theory Models

✱Jian Zhu, The University of Michigan, 1420 Washington Heights, Department of Biostatistics, Ann Arbor, MI 48109, *jianzhu@umich.edu*; Trivellore Raghunathan, The University of Michigan; Raymond Bingham, The University of Michigan

Key Words: multiple imputation, scale variables, item response theory, Gibbs sampling

Creating scale variables when some items are missing is a common but difficult problem. This paper uses item response theory models as a vehicle for imputing the missing items that incorporates scale-item response structure in the imputation process. We use a Bayesian framework and normal ogive IRT models. Approximately independent draws are obtained from the joint posterior distribution of item parameters, latent scales and missing item responses through Gibbs sampling. We use a longitudinal study with numerous scale-items to illustrate the methodology and contrast it with other standard imputation methods. We also report on a simulation study to assess the repeated sampling properties of various population level estimates derived using this methodology.

A Comparison of Multiple and Single Imputation Algorithms for Missing Data in Quadratic Discriminant Function

✱Page C. Moore, University of Arkansas for Medical Sciences, 4301 W. Markham St., Slot #781, Little Rock, AR 72205, *pmoore@uams.edu*; Songthip Ounpraseuth, University fo Arkansas for Medical Sciences; Dean Young, Baylor University

Key Words: Quadratic Discriminant Function, Statistical Classification, Single Imputation, Multiple Imputation, Missing Data, MAR

Because researchers are increasingly using multivariate statistical methods in research areas such as medicine, genome sequences modeling and clinical trials, statisticians are often faced with the dilemma of dealing with data sets containing missing observations. The presence of incomplete observations is increasingly important in statistical analysis as the number of variables measured increases. We examine the performance of several commonly-used methods of handling missing values, such as mean substitution as well as selected multiple imputation (MI) methods. Specifically, we evaluate the performance of the Quadratic Discriminant function after replacing missing training observations with data imputed under various methods. We use a Monte Carlo simulation to examine the expected probabilities of misclassification under each of these single and multiple imputation methods.

Imputation of Nondetects in Environmental Concentration Measurements

✱John Rogers, Westat, 1650 Research Blvd., Rockville, MD 20850, *JohnRogers@westat.com*

Key Words: Detection limit, Imputation, Non-detect, Environmental

Nondetects are left-censored measurements generally reported by the laboratory as less than a specified detection limit. Analysis of data with nondetects typically involves substituting an assumed value for the nondetects, such as half the detection limit, or using models that explicitly handle censored data such as survival models using an assumed distribution for the observations. Substitution of an assumed value is a form of imputation. The usefulness of the imputation can be improved by using correlated data. This paper presents several approaches to imputation of nondetected values based on an assumed distribution for the data and on other correlated data. Examples of correlated data include related but uncensored measurements and the preliminary concentrations used to determine if the reported concentrations are above or below the detection limit.

Exact and REML-Based Confidence Intervals for Variance Components Using Non-normal Distributions

✱Brent Burch, Northern Arizona University, Department of Mathematics & Statisitics, PO Box 5717, Flagstaff, AZ 86011-5717, *brent.burch@nau.edu*

Key Words: linear mixed models, heritability, variance components, REML, confidence intervals

Using normality assumptions, one can compute exact confidence intervals for variance components. Exact intervals differ from intervals that are based on the asymptotic distribution of restricted maximum likelihood (REML) estimators. Simulation studies are conducted to evaluate the performance of the exact and the REML-based intervals when the observations are not normally distributed. Simulated coverage probabilities and expected lengths provide guidance as to which interval procedure is favored for a particular scenario. Confidence intervals for heritability, a particular function of variance components, are also considered.

Restricted Heteroscedastic IV Estimators: A Simulation with Small Samples

✱Luis Frank, University of Buenos Aires, Av. San Martin 4453, College of Agronomy, Buenos Aires, C1417DSE Argentina, *lfrank@agro.uba.ar*

Key Words: instrumental variables, IV estimator, restricted models, linear models, regression analysis, small samples

Given y=Xfl+e subject to the Rfl=r, the paper proposes four different instrumental variables estimators to handle not "well behaved" data, and simulates them with small samples. A monotone transformation of the columns of X is used to define a matrix Z of instruments that set to zero log transformed elements when x(ij) tend to 0, situation which arises frequently under a Cobb-Douglas demand function specification. Preliminary results show fairly good performance (in terms of unbiasedness and efficiency) of the restricted instrumental variables approach despite some numerical drawbacks presented by one of the estimators with extremely small samples.

Covariate Measurement Error Correction via Empirical Bayesian Methods

✱Miguel Padilla, Old Dominion University, 1665 University Blvd., RPHB 327, Birmingham, AL 35233, *mpadilla@uab.edu*; Daniel Shriner, The University of Alabama at Birmingham

Key Words: measurement error, linear model, Empirical Bayes, reliability

Statistical linear models are widely used to analyze a variety of data from the behavioral to the biological sciences. However, all variables in a linear model are assumed to be measured without error. Here, we investigate measurement error in a covariate. Measurement error in a covariate tends to attenuate its associated parameter estimate and confound the residual variance. We present an Empirical Bayesian method for correcting measurement error problems in linear models with emphasis on correcting a measurement error-contaminated covariate rather than just correcting parameter estimates. The method is assessed by computer simulation in terms of Type I error and power. Factors investigated are degree of measurement error and sample size. Results indicate that our method can be used to correct covariate measurement error in linear models.

A Single-Step Modified Bonferroni Procedure for Multiple Tests

✱Michelle A. Roozeboom, Iowa Foundation for Medical Care, West Des Moines, IA 50266, *roozeboommic@hotmail.com*; Daniel Mundfrom, University of Northern Colorado; Jamis Perrett, Texas A&M University

Key Words: Regression, Type I Error, Statistical Power, Simulation, Multiple Tests, Bonferroni Adjustment

The Bonferroni method maintains a Type I error rate less than or equal to the nominal level when multiple tests are performed. However, this adjustment reduces power. The purpose of this research was to establish a single-step method for multiple tests that controls Type I error while increasing power. Sample data (n = 25, 50, 200) were simulated to have normal distributions for hypothesis testing of 3 to 20 beta coefficients. Correlations between x and y and among the xs was varied. A new denominator related to the number of tests performed was less than or equal to k. Mean Type I error for Bonferroni, Holm, New, and No Adjustment were .031, .035 , .055, and .317 (SDs= .011, .011, .019, .115, and N=2958). Mean statistical power was .339, .344, .361, and .496 (SDs=.391, .394, .390, .363). The new method resulted in increased statistical power while correcting for inflated Type I error.

Regression Analysis for Line Scale Data in Sensory Science

✱Lin Xie, Kansas State University, 1541 International Ct, N12, Departmnet of Statistics, Manhattan, KS 66502, *lxie@ksu.edu*; Rui Xiong, Unilever Home & Personal Care

Key Words: Regression analysis, line scale data, preference mapping

Line scales are widely used in sensory evaluation and consumer testing. The scores of response variables from line scales take only values within a specified range (e.g., 0-10, 0-100). In preference mapping and sensory-instrumental relationship, the response variables are often linearly regressed on some other explanatory variables. Since regression models are inherently unbounded (-8, +8), they could produce predicted values that are outside the ranges of the line scales for some response variables. Conventionally, if predicted values are outside the specified range, they are set to either the minimum or maximum value of the scales. In the paper, an alternative approach, which transforms the line scale range to unbounded range (-8, +8) before regression analysis, is proposed to address the problem.

Model-Building for Mixed-Effects Models with Random Scale Effects

✱Lei Shu, Purdue University, 250 N. University Street, Statistics Department

250 N.University Street, West Layfayette, IN 47907-2066, *lshu@stat.purdue.edu*; William Cleveland, Purdue University

Key Words: Mixed-effects model, Model building, Random scale effects, Data visualization

The use of mixed-effects models in practice, often in the form of Bayesian hierarchical models, is growing rapidly because of major developments in computational methods for these models. Mixed-effects models with measurements of a response on a continuous measurement scale typically specify the random effects as location effects. But the data sets with random location effects can also have random scale effects. Typically, random location effects are taken to be normal and random scale effects to be square root inverse gamma. But in practice, other distributions often occur. This presentation introduces a sequential model building approach for the mixed-effects models with random scale effects. Data visualization methods play a fundamental role in all phases of this model building process: data exploration, model identification, and model-checking.

Modeling Generalized Linear Mixed Models in S-PLUS, R, and SAS

✱Stella Karuri, Insightful Corporation, 1700 Westlake Avenue, Suite 500, Seattle, WA 98109, *swkaruri@insightful.com*; Tim Hesterberg, Google, Inc; Chris Fraley, Insightful Corporation

Key Words: generalized mixed models, laplace estimation, penalized quasi likelihood, maximum likelihood

Generalized Linear Mixed Models (GLMMS) are popular in modeling clustered data. We survey existing routines for fitting GLMMS implemented in R, S-PLUS and SAS, and find that all routines fail in some situations. We describe a new S-PLUS routine, based on Spherical Radial Quadrature that offers more reliable estimation.

Measuring Goodness-of-Fit for Generalized Linear Models

✱Yves Chretien, Harvard University, Department of Statistics, Cambridge, MA 02139, *chretien@stat.harvard.edu*; Carl Morris, Harvard University

Key Words: GLMs, goodness-of-fit, residuals, diagnostics, information theory, exponential families

Linear regression enjoys a well-known array of diagnostics for evaluating model fit, such as residual plots and coefficients of determination. Generalized linear models (GLMs) possess no such "standard" suite of diagnostic techniques. Those techniques that do exist are more ad hoc and do not fit as readily into a unified framework as for linear regression. Information theory provides a basis for generalizing measures of model fit and residual variation to GLMs. We demonstrate applications in logistic regression, using both simulation studies and data from an ongoing national survey analysis project. While many generalizations of the linear regression toolkit are possible, information theory combined with the special properties of exponential families supports a natural and readily interpretable framework for extending the familiar linear regression diagnostic suite to GLMs.

Comparison of Generalized Linear Models To Evaluate Factors Associated with Metabolic Syndrome

✱Desta B. Fekedulegn, National Institute for Occupational Safety and Health, National Institute for Occupational Safety and Hea, 1095 Willowdale RD, Morgantown, WV 26505, *djf7@cdc.gov*; Michael E. Andrew, National Institute for Occupational Safety and Health; John M. Violanti, The State University of New York at Buffalo; Tara A. Hartley, National Institute for Occupational Safety and Health; Luenda E. Charles, National Institute for Occupational Safety and Health; Cecil M. Burchfiel, National Institute for Occupational Safety and Health

Key Words: Metabolic syndrome, generalized linear models, count data

Our goal was to evaluate potential use of several models where the outcome represents a count, rather than presence or absence, and estimate measures of association. Metabolic syndrome defined by Grundy (2005) is a binary outcome that refers to presence of three or more of five components, but components can also be summed as a count. Data from 96 randomly selected Buffalo police officers were used to examine associations between depression score (CES-D) and metabolic syndrome. Measures of association for 1 SD change in CES-D were calculated using both the binary outcome and count of syndrome components. Poisson regression used to estimate ratios of means for associations with count as an outcome were statistically significant yet methods using binary outcomes were not. This example shows the potential for Poisson regression to provide an interpretable alternative measure of association.

Joint Distribution of Parametric and Nonparametric Parameters in Longitudinal Factorial Designs

✱Terri Wooten, University of Arkansas for Medical Sciences, *terri.wooten@arkansas.gov*; D. Keith Williams, University of Arkansas for Medical Sciences; Zoran Bursac, University of Arkansas for Medical Sciences

Key Words: Longitudinal, nonparametric, parametric assumption departure

We present preliminary work involving the estimation of the joint distribution of parametric and nonparametric parameters on the same data set. Estimates of the parameters and their corresponding covariance structures enable one to determine if the outcomes are the same, utilizing the joint distribution of these parameters. The result is consistent with the idea of inter-rater reliability. If parameters agree then perhaps assumption departures or outliers are not highly problematic. If the reverse is true, then further data review is necessary. In this work we derive a departure test statistic, its approximate distribution, and a hypothesis test. Additionally, simulation results are presented to estimate the sensitivity properties of the departure statistic.

Correlated Binary Data: A Comparison of Analytic Approaches for the Analysis of Patient-Reported Outcomes in Diarrhea-Predominant Irritable Bowel Syndrome

✱Norman Bohidar, Johnson & Johnson Pharmaceutical R&D, LLC, 445 Belle Lane, Harleysville, PA 19438, *norman.bohidar@comcast.net*; Mary Ellen Frustaci, Johnson & Johnson Pharmaceutical R&D, LLC; Surya Mohanty, Johnson & Johnson Pharmaceutical R&D, LLC

Key Words: Binary Data, Correlated Data, Patient Reported Outcome, GEE

The patient reported outcome of daily control of urgency is commonly collected in the study of diarrhea-predominant irritable bowel syndrome (d-IBS). Published analyses of this endpoint focus on the median of the percentage of days that adequate relief is reported, typically over 12 weeks of treatment. Alternatively the data can be analyzed as correlated binary data using a Generalized Estimating Equations approach. This presentation will use published data to determine the range of expected correlation and treatment effect in order to compare the power of this alternative approach to the standard. In addition, the comparison of having 4 weeks of data versus 12 weeks of data will be evaluated.

GEE Performance in Analyzing Longitudinal Binary Data Collected Over Unequal Time Intervals

✱Soe Soe Thwin, Boston University Medical Center, Geriatrics Section, Robinson-2, 88 East Newton, Boston, MA 02118, *sst@bu.edu*; David Gagnon, Boston University School of Public Health; Howard Cabral, Boston University School of Public Health; Adrienne Cupples, Boston University School of Public Health

Key Words: Longitudinal Data, Unequal Intervals, Binary Response, Estimation, GEE, ALR

Binary response data observed over unequal time intervals were generated with a modified AR(1) correlation structure, whereby, pair-wise correlations were adjusted based on their interval length relative to the interval between the first 2 observations. This structure simplifies to the AR1 structure when time intervals are equal. Bias in parameter estimation and relative efficiency were evaluated for scenarios when working correlation structures were misspecified or when alternating logistic regression was invoked. We observed that performance of GEE depends on characteristics of the study design, namely the sample size (Large vs. Small), covariate (Time Stationary vs. Time Varying), and cluster Size (# observations per subject). Adjustment of interval length in correlation structure provided better estimate for highly correlated data, especially for time-stationary covariates.

Capturing Group Membership via Growth Mixture Models: A Simulation Study

✱Kevin Delucchi, University of California, San Francisco, 401 Parnassus Ave., Box 0984-TRC, San Francisco, CA 94143, *kdelucchi@lppi.ucsf.edu*; Alan Bostrom, University of California, San Francisco

Key Words: latent variable models, longitudinal, growth mixture, simulation

Results of a series of simulations examining the ability of latent growth mixture models to capture population group membership, intercept and slope parameters is presented. The model was a common design; longitudinal with four equally-spaced assessments and two latent groups. 56 conditions were simulated from a 7 x 2 x 2 x 2 factorial design with 1000 samples per condition for total Ns of 300 and 900 with sample size imbalance ranging from small to large. Percent correctly classified ranged from 58% to 88%. Estimates of population slope and intercept ranged from acceptable to poor depending on conditions simulated. Additional simulations to be reported indicate large sample sizes are needed for this type of modeling. The addition of an informative covariate improved the estimates.

Pseudo Likelihoods for Nonlinear Regression Models

✱Harshini Fernando, Purdue University North Central, Dept. Mathematics and Statistics, 1401 South US Highway 421, Westville, IN 46391, *hfernando@pnc.edu*

Key Words: Pseudo likelihoods, conditionally linear regression models, elliptically contoured models

We consider robust methods of likelihood inference for the nonlinear parameter in a conditionally linear nonlinear regression models. A closed-form expressions for robust conditional, marginal, profile and modified profile likelihood functions for the nonlinear parameter under elliptically contoured data distributions are derived. Next, several well-known examples are considered and Monte Carlo simulation results are presented.

On Interval Estimations for a Dichotomized Variable Derived from a Continuous Measure

✱Jessica L. Kim, U.S. Food and Drug Administration, 1401 Rockville Pike, Rockville, MD 20852, *jessica.kim@fda.hhs.gov*; Jean Wang, U.S. Food and Drug Administration

Key Words: dichotomized variable, tolerance limit, data simulation, RBC (Red Blood Cell) products, percent of coverage

When an analysis is done with a dichotomized variable from a continuous measurement, the result can be interpreted differently. In this paper, tolerance limits with various coverage levels using the original data scale will be estimated. Then, the results of tolerance intervals will be used as a threshold value to dichotomize the variable. The corresponding tolerance limits of the dichotomized variable will be investigated and compared to the results of the original scale. Data from simulation as well as in vivo studies of Red Blood Cell products will be discussed and commented as an application. An interval that predicts p percent of coverage for a dichotomized variable derived from a continuous measure will be investigated.

Shrinkage Drift Parameters Estimation for Multifactors Ornstein-Uhlenbeck Processes

✱Ejaz S. Ahmed, University of Windsor, 401 Sunset Avenue, Windsor, ON N9B 3P4 Canada, *seahmed@uwindsor.ca*; Severien Nkurunziza, University of Windsor

Key Words: Ornstein-Uhlenbeck process, Vasicek process, Gaussian process, Gaussian process, shrinkage estimator, MLE

We consider some inference problem concerning the drift parameters of multi-factors Vasicek model (or multivariate Ornstein-Uhlebeck process). For example, it asserts that the term structure of interest rate is not just a single process, but rather a superpositions of several analogous processes. This motivates us to develop an improved estimation theory for the drift parameters when homogeneity of several parameters may hold. However, the information regarding the equality of the theses parameters may be imprecise. In this context, we consider {\it Stein-rule} (or shrinkage) estimators which allow us to improve upon the performance of classical maximum likelihood estimator (MLE). Asymptotic properties of some shrinkage estimators are studied. Under an asymptotic distributional quadratic risk criterion, their relative dominance picture is explored and assessed.

A Simple Approximation to the Likelihood Interval for a Binomial Proportion

✱Ian Harris, Southern Methodist University, Department of Statistical Science, PO Box 750332, Dallas, TX 75275-0332, *iharris@mail.smu.edu*

Key Words: quadratic log likelihood, Wald interval, parameter transformation

A likelihood interval for the probability of success p, in a binomial distribution is formed from inversion of the log likelihood ratio. The endpoints of the interval must be found using numerical methods, so the interval is not suitable for use by practitioners who are looking for quick or “back of the envelope” calculations. We propose a simple, accurate, closed form approximation to the likelihood interval based on a recentered form of the Wald interval.

Bivariate and Multivariate Association Coefficients

✱Nan Kong, Educational Testing Services, 270 Hampshire Dr., Plainsboro, NJ 08536, *nkong@ets.org*

Key Words: Entropy, Conditional entropy, L-bivariate association coefficient, L-multivariate association coefficient, K-dependence coefficient

Given a system of multiple random variables, a new measure, which is called L-multivariate association coefficient, is defined using (conditional) entropy. Unlike traditional correlation measures, the L-multivariate association coefficient measures the degree of the associations among the multiple variables in the given system. The L-multivariate association coefficient for the system of two random variables is also called the L-bivariate association coefficient. The L-multivariate association coefficient is a symmetric measure which measures both linear and nonlinear associations among the variables in the system. A direct application of the L-multivariate association coefficient is variables selection or variables reduction.

Heterogeneous Variance Models and Their Applications in Parameter Designs

✱Fassil Nebebe, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, QC H3G 1M8 Canada, *fassil.nebebe@concordia.ca*

Key Words: Control factors identification, crossed array designs, PerMIA, model mis-specification

We consider in this paper the issues of estimation and variance model building in the analysis of robust parameter designs. The model examined accommodates the possible simplification resulting from the use of a “Performance Measure Independent of Adjustment (PerMIA)” and is applicable in both crossed array and combined array designs. We consider a likelihood based parametric method through the use of a very flexible class of distributions for the error terms. The application of the idea of PerMIA typically involves a smaller set of “control factors” that affect the PerMIA (the remaining design factors are adjustment factors that affect only the mean). Thus a central issue is the identification of the control factors. We address the issue of identifying the control factors in building an equation for the PerMIA, and examine the effects of model misspecification.

Residual Analysis for Detecting Mismodeling in fMRI

✱Ji Meng Loh, Columbia University, Dept of Statistics,, *meng@stat.columbia.edu*

Key Words: fMRI, residual analysis, mismodeling

Functional magnetic resonance imaging (fMRI) data is often analyzed via the voxel-wise general linear model (GLM) approach. A relatively small amount of mis-modeling, however, can result in severe power loss, and inflate the false positive rate. We propose a simple procedure that can be used to identify possible voxels or regions of the brain where model misfit may be present. The key idea is that if there is model misfit, residuals will be systematically larger in mismodeled segments of the time series. By looking at the weighted sum of consecutive residuals using a moving window, our method can pick out regions of a residual time series in which the residuals are consistently larger than expected by chance, while ignoring spurious large residuals. The effectiveness of the procedure is examined using a simulation study and by applying to an fMRI data set.

Receiver Operating Curves (ROC) in the Evaluation of Diagnostic Accuracy in Medical Devices

✱Inder Sharma, Sharma Associates, Inc., 11012 Persimmon Creek Dr., Mint Hill, NC 28227, *sharmaasoc@carolina.rr.com*

Key Words: ROC, Sensitivity, Specificity, False-Positive, True-Positive, Accuracy

ROC curve analysis is an evaluation technique used in the past extensively in signal detection, machine learning and information retrieval. In medical devices traditionally an accuracy of a test has been relied on measures such as sensitivity and specificity of a test. But sensitivity and specificity depend upon the cut point used to define the disease status. For certain diseases it is difficult to make a clear distinction of disease status due to the subjectivity involved and one has to depend upon scales that go beyond disease or no disease classification. ROC Curve provides a single measure to provide a measure of accuracy of a test. ROC curves plot false-positive rates on the x-axis and true-positive rate on the y-axis. The construction of ROC Curves and it's advantages as well as limitations for measuring accuracy of tests for the above situations will be presented.

Assay Performance Comparison

Victoria Petrides, Abbott Laboratories; ✱Susan Gawel, Abbott Laboratories, 100 Abbott Park Rd., Building AP20-2, Dept. 09B6, Abbott Park, IL 60064, *susan.gawel@abbott.com*; Anthony Orzechowski, Abbott Laboratories; Theresa Garwood, Abbott Laboratories

Key Words: Total, Error, Assay, Diagnostic, Device, Sigma

Labs are faced with the challenge of evaluating the performance of an assay on a given platform based on information provided by manufacturers, usually in the form of product labeling. The type and format of information provided by each manufacturer varies, making it difficult for labs to compare the performance of similarly available assays. The objective of this poster is to show multiple methods for evaluating the performance of an assay and highlighting the method that seems most comprehensive. Bias and imprecision were determined on multiple platforms for a given assay and were used in conjunction with the total allowable error to calculate a single performance metric—the sigma value—making it easier to evaluate the performance of the assay across platforms.

Methods for Comparing Biomarkers for Aid-in-Diagnosis and Prediction

✱Sarah Hurwicz Kogut, i3 Statprobe, 12125 Technology Drive, Eden Prairie, MN 55344, *sarah.kogut@i3statprobe.com*

Key Words: diagnostic test, biomarker test, receiver operating characteristic curve, ROC curve, AUC

This poster will display examples of diagnostic tests using several biomarkers for aid-in-diagnosis. Use of several medical analysis software packages for comparison of ROC curves will be demonstrated. The study hypothesis should govern choice of analysis method (Cook, 2007; Pepe, Longton, and Janes, 2008). Multiple biomarker test measurements taken on the same patient can result in correlated receiver operating characteristic (ROC) curves. Comparison of correlated ROC curves may require alternate methods than are used in commonly available software (for example, Hanley and McNeil, 1982). Correlated data collected under a paired design can be analyzed using the DeLong et al (1988) method. It may be that only one part of the ROC curve is of interest in the use of a particular biomarker test. In this case, the test comparing partial AUCs may be more appropriate (McClish, 1989).

Importance in the Developing Countries of Statistics in Sector Security and Defence Policy

✱Jhoner L. Perdomo, V, Universidad Central de Venezuela, Student Center Research EECA, School Statistics, Resid 3, Caracas, International 1010 Venezuela, *elucevista@hotmail.com*

Key Words: Statistics Military, Statistical Security, Peace and Defense

The approach in this research arises because of the need to have these tools in developing countries and the historical background documentaries serving of arguments to highlight the importance of its application in the field of public security and national defence. It is a work of historical research as a that documentary tells as the statistic has contributed and has been implemented for several historical processes relevant to mankind. It discusses among many, the two most important cases of the great strategists Sun Tzu and Clausewitz, highlighting their usefulness and including statistics in his classic books of War, The Art of War, and The War, respectively.

Confidential Data Perturbation via Skew-T Distributions

✱Seokho Lee, Texas A&M University, Department of Statistics, 3143 TAMU, College Station, TX 77843-3143, *leesh@stat.tamu.edu*; Marc G. Genton, University of Geneva; Reinaldo B. Arellano-Valle, Pontifícia Universidad Católica de Chile

Key Words: Confidentiality, Database Management, Kurtosis, Quantile, Security, Skewness

The protection of confidential variables in governmental, commercial and medical databases has become a very important problem. To this end, a new data perturbation method (STDP) for database security based on skew-t distributions is proposed. Unlike normal distributions, the more general class of skew-t distributions is a flexible parametric family that can model skewness and heavy tails in the data. Because databases having a normal law are seldom encountered in practice, this newly proposed method is of great interest for database managers. The performance of STDP is investigated by a Monte Carlo simulation study and its superiority is demonstrated. Of particular importance is the ability of STDP to accurately reproduce characteristics of the tails of the distribution, such as queries based on upper quantiles. The STDP method is applied to a medical database related to breast cancer.

406 Section on Health Policy Statistics Speaker with Lunch (fee event)

Section on Health Policy Statistics
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Weight and Mortality: Estimates for the U.S. Population

✱Katherine M. Flegal, Centers for Disease Control and Prevention, 3311 Toledo Rd,, Hyattsville, MD 20782, *KFlegal@cdc.gov*

Key Words: obesity, mortality, attributable fraction, sample surveys, NHANES

Estimates of the population burden of all-cause and cause-specific mortality associated with different levels of body weight have been derived by combining data from three sources: the relative risks of mortality associated with different weight levels based on follow-up of three waves of nationally representative survey data collected from 1976–1980; the more current distribution of body weight and other covariates derived from more recent national survey data; and national vital statistics data on numbers and causes of death in the U.S. population. The methodology and results from these investigations will be described, as well as some of the issues that have arisen in attempting to communicate the findings to technical and nontechnical audiences.

407 Section on Bayesian Statistical Science Roundtables with Lunch (fee event)

Section on Bayesian Statistical Science
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Bayesian Methods in Medical Device Trials

✱Bradley P. Carlin, The University of Minnesota, Division of Biostatistics MMC 303, 420 Delaware St. S.E., Minneapolis, MN 55455, *brad@biostat.umn.edu*

Key Words: Bayes, adaptive clinical trials, type I error, power, FDA

Bayesian methods are increasingly popular in medical device trials, as evidenced by the recently released FDA Guidance Document. Bayesian methods offer several key advantages, including easy incorporation of information from earlier studies, the ability to borrow strength across studies or centers, and a framework well-suited to adaptation (say, of the patient allocation ratio) that results in more ethical treatment of patients. However, such trials retain a strong frequentist flavor, due largely to the continued insistence by FDA that all trials have good frequentist operating characteristics (power and Type I error). We will discuss this situation and attempt to find a middle ground between frequentist long-run protection and Bayesian flexibility.

Bayesian Methods in the Evaluation of Modalities for Diagnosis and Prediction

✱Constantine Gatsonis, Brown University, Center for Statistical Sciences, Box G-S121-7, Providence, RI 02912, *gatsonis@stat.brown.edu*

Key Words: diagnosis, prediction, biomarkers, ROC analysis, sensitivity, specificity

The Bayesian literature on the design and analysis of studies of diagnostic tests and biomarkers has had considerable growth in recent years. The majority of published work proposed parametric and nonparametric methods for studies assessing the sensitivity and specificity and/or the ROC curve of tests. Relatively less attention has been paid to the assessment of the predictive value of tests and questions of study design. In this roundtable, we will discuss current developments in this area of Bayesian methods and examine the experience from applying such methods in practice.

408 Biopharmaceutical Section Roundtables with Lunch (fee event)

Biopharmaceutical Section
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

A Bayesian Approach to the ICH Q8 Definition of Design Space

✱John J. Peterson, GlaxoSmithKline Pharmaceuticals, Research Statistics Unit (UP-4315), 1250 So. Collegeville Road, Collegeville, PA 19426, *john.peterson@gsk.com*

Key Words: Gibbs sampling, noise variables, quality by design, response surface methodology, robust parameter design, six sigma

A key issue for the pharmaceutical industry and the FDA is the ICH Q8 concept of "Design Space," as applied to pharmaceutical and biopharmaceutical manufacturing. Design Space is essentially a set process capability operating conditions. This roundtable will discuss the ICH Q8 definition and a Bayesian formulation for it.

Do Statisticians Still Look at the Data?

✱Bret J. Musser, Merck Research Laboratories, 126 E Lincoln Ave, Rahway, NJ 07065, *bret_musser@merck.com*

Key Words: industry, exploratory data analysis, visualization, software

Statisticians working in the pharmaceutical industry must often put in extra effort to keep in touch with the data. The complexity of new standardized data structures, the need to produce publication-ready tables, and the demise of truly interactive data analysis tools are pushing statisticians to stop looking at the data. Are today's graduates being trained to look at data in the face of these challenges?

CANCELLED - Prespecifying Analyses: In the Lifecycle of a Clinical Trial, Does It Matter When, Where, and How You Prespecify?

Key Words: pre-specification, post-hoc, analysis plan

A critical criterion to the integrity of any statistical analysis from the regulatory and scientific point of view is whether it was "pre-specified," as opposed to "post-hoc." The distinction often lies in when the analysis was performed versus the analyst's state of knowledge of the data. This roundtable will discuss whether pre-specification is a simple yes/no issue, or more a continuum. Participants should be prepared to contribute with opinions and experiences on such related topics as the importance of how and where such pre-specification is done and whether the audience matters.

Patient-Reported Outcomes: Measurement Reliability and Sensitivity

✱Tammy Massie, U.S. Food and Drug Administration, 1401 Rockville Pike, Rockville, MD 20852, *Tammy.Massie@fda.hhs.gov*

Key Words: Patient Reported Outcomes, Categorical Response, Clinical Trial

A PRO (patient-reported outcome) is a measurement of any aspect of a patient's health that comes directly from a patient. In the usual controlled clinical trials, PROs can be considered the primary, secondary, or tertiary endpoints. Within PROs, subjects may provide responses including continuous, categorical, and open-ended responses over time. Frequently, PROs contain repeated measures categorical response data from each subject. These PROs will be the subject of discussion. Data collected may be analyzed using statistically valid categorical data analysis methodologies. The proper design and use of PROs present considerable challenge to statisticians. Using the PRO guidance to stimulate dialogue, we will explore issues related to the reliability, sensitivity, and analysis of categorical response PROs in clinical trials.

Applying Lean Six Sigma Concepts to Adaptive Trial Logistics

✱Darcy Hille, Merck & Co., Inc., Mailstop UG1C-46, 351 Sumneytown Pike, North Wales, PA 19454, *darcy_hille@merck.com*; Paulette Ceesay, Merck & Co., Inc.

Key Words: adaptive clinical trials, Lean Six Sigma

Adaptive trials are clinical trials that use data not available at study start as a basis for modifications to the trial design. Lean Six Sigma is a business improvement methodology that stresses both efficiency and quality. Anticipating the complex logistics necessary for efficient, repeatable, adaptive trials, we are applying concepts from Lean Six Sigma to implement an adaptive randomization program at Merck. We hope to engage in a discussion about the challenges and successes involved in the application of Lean Six Sigma to the adaptive trial process.

409 Section on Government Statistics Roundtable with Lunch (fee event)

Section on Government Statistics
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Cell Phone Survey Sampling and Weighting

✱Michael P. Battaglia, Abt Associates, Inc., 55 Wheeler Street, Arlington, MA 02138, *mike_battaglia@abtassoc.com*

Key Words: Dual Frame Sample Designs, Cell Phone-Only Households, Sampling Weights, Cellular Sampling Frame, Nonreponse, Ported Numbers

Landline random-digit-dialing survey coverage problems due to cell phone–only households have led to a growing number of cellular telephone surveys being conducted to supplement traditional landline random-digit-dialing telephone surveys. We will discuss cell phone sampling frames; landline to cell ported numbers; trade-offs of screening for cell phone–only households versus using a dual-frame sample design that includes households with landline and cellular telephone service; the two alternative sampling models that have received the most attention—a household level sampling model versus a person level sampling model; and weighting techniques, including how the National Health Interview Survey, Current Population Survey, American Community Survey, and others can be used to develop estimates of the size of the cell phone–only population.

410 Section on Physical and Engineering Sciences Roundtable with Lunch (fee event)

Section on Physical and Engineering Sciences
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

How Quality Evaluation Affects Product Reliability Models

✱Dan Fitzsimmons, Boeing Commercial Airplanes, PO Box 3707, MC 08 51, Seattle, WA 98124-2207, *dan.k.fitzsimmons@boeing.com*

Key Words: Sampling, Consumer Risk, Accuracy, Measurement Uncertainty, Reliability, Initial Reliability Requirement

In some industries, when nonconforming product is discovered in a product stream, it is replaced with product that has been carefully inspected, and manufacturing personnel are at least as careful with product that will be inspected as with product that will not. In those cases, classical mathematical models of outgoing quality say a product stream that is inspected more will produce units of product with a higher probability of conformance. That higher probability of conformance may have an effect on the product's reliability budget. If so, then high product reliability requirements may place mathematical constraints on the application of sampling and/or the accuracy of product evaluation tools and methods. This roundtable is an opportunity to discuss cases where this connection may have significant applications.

411 Section on Statistical Consulting Roundtable with Lunch (fee event)

Section on Statistical Consulting
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Is That Clear? Communicating with Clients Effectively

✱Novie O.M. Younger, The University of the West Indies, Jamaica, P.O. Box 25580, TMRI, Kin # 3960, Miami, FL 33102, *novie.younger@uwimona.edu.jm*

Key Words: statistical jargon, communicating effectively, policy, use of statistics in daily life

Heavy use of statistical jargon can be a deterrent to the application of statistical inference or appropriate use of results. As statisticians, the make-up of our clientele ranges from students who attend our classes to collaborators in research to media personalities and lay persons trying to make sense of the results we produce. How can we ensure our results move beyond high-impact, peer-reviewed journals to active use in industry and everyday life? It begins with making our results more lucid to persons with little exposure to statistics. Our discussion will explore methods of effectively communicating results and use of methodology to persons who need to use the statistical tools and apply results of research but are without a strong numerate background.

412 Section on Statistical Education Roundtables with Lunch (fee event)

Section on Statistical Education
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Statistical Outreach: How Statisticians in Industry and Academia Can Reach Out to Secondary Educators and High-School Students

✱Cynthia A. Gargano, Merck & Co., Inc, UG1D-44, P.O. Box 1000, North Wales, PA 19454, *cynthia_gargano@merck.com*

Key Words: outreach, AP statistics, students, high school, teachers

As statisticians in academia and industry, it is our duty to inform teachers and students about exciting opportunities in the statistics field. In addition, high-school statistics teachers welcome the assistance and consultation of statistics professionals. High-school statistics teachers face many challenges, including how to completely cover AP exam material and how to convince college entrance boards that AP Statistics is as valuable, if not more so, than calculus. However, it is often a challenge to coordinate and focus these outreach efforts. Academia and industry statisticians face difficulties coordinating outreach meeting times, discussion topics, and narrowing the focus of the outreach. The purpose of this roundtable will be for people to share ideas about outreach: what has worked in the past, what has not worked in the past, and what may work in the future.

Active Learning: Engaging Students in Learning Now

✱Paul J. Fields, Brigham Young University, 223A TMCB, Provo, UT 84602, *pjfields@byu.edu*

Key Words: Active-Learning, Engagement, Learning Outcomes

It is commonly accepted among educators that active learning is more effective than passive learning. However, students are often resistant to fully engaging in active learning. This seems to be an example of the old adage, “You can lead a horse to water, but you cannot make him drink.” If active learning is better for students, why do many of them not “drink”? We will examine possible explanations ranging from student apathy to lack of skills requisite for active learning. We will discuss ideas for helping students overcome their resistance to full engagement in active learning so they can achieve maximum personal learning outcomes. Further, we will identify resources available within the statistic education community to facilitate student “buy in” to active learning. We also will form a collaborative effort to exchange techniques to enhance student engagement in learning now.

Broadening the Undergraduate Statistics Curriculum

✱Deborah Nolan, University of California, Berkeley, Dept of Statistics, 367 Evans Hall MC 3860, Berkeley, CA 94720-3860, *nolan@stat.berkeley.edu*

Key Words: Education, Computing

The statistics curriculum has not kept pace with the practice of statistics. We will discuss ways to design syllabi, starting from scratch, making use of the large collection of computational tools and statistical experiences at our disposal. We advocate embracing computing as an essential building block of statistical creativity and practice. If our students are facile with computing, they will be in a position to work on interesting topical scientific problems.

413 Section on Statisticians in Defense and National Security Roundtable with Lunch (fee event)

Section on Statisticians in Defense and National Security
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Research Directions in Information Integration

✱Wendy Martinez, Office of Naval Research, 805 Shaw Court, Fredericksburg, VA 22405, *wendy.martinez@navy.mil*

Key Words: Data mining, Inference, Data analysis, Knowledge discovery

There is an increasing need in defense and national security for the automated integration of disparate sources and types of data. One would like to use the data or information that results from the integration process in inference, prediction, synthesis, data mining, and so forth. The expectation is that an analyst might gain better performance in these tasks through the fusion and integration of these disparate sources of data. The purpose of this roundtable is to discuss new research directions in information integration with an emphasis on how statisticians can contribute to this important area.

414 Section on Statistics and Marketing Roundtable with Lunch (fee event)

Section on Statistics and Marketing
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Applications of Internet Clickstream Data: Past, Present, and Future

✱Wendy Moe, University of Maryland, 20742, *wmoe@rhsmith.umd.edu*

Key Words: clickstream data, internet, e-commerce, marketing

Early internet research that used clickstream data was primarily descriptive in nature. The goal was simply to understand how consumers behaved in the new e-commerce environment. Research questions focused on modeling the processes behind store visitation, purchase conversion, and search behavior (both across and within store sites). These studies provided frameworks with which to address future e-commerce research problems. Now, the focus has shifted from the descriptive to the prescriptive, using internet data to predict consumer responses to market events, forecast product sales, and identify key drivers of purchasing. These applications will be the focus of our roundtable discussion.

415 Section on Statistics in Epidemiology Roundtable with Lunch (fee event)

Section on Statistics in Epidemiology
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Statistical Innovation in Submissions to FDA and Other Regulatory Bodies

✱Diane Fairclough, University of Colorado Denver, 2570 S Jackson St., Denver, CO 80210, *dianefairclough@earthlink.net*

The goal of this roundtable is to discuss strategies that have been/could be successful or unsuccessful for incorporating statistical innovation into applications for drug and device applications to the FDA and other regulatory bodies. For example, one area where there has been considerable methods research is in the handing of missing data. What strategies have been successful in changing the primary data analysis from complete case or change from baseline to last observation, to those that include all available data, or use methods such as joint (shared parameter) models, pattern mixture models, or multiple imputations? Other examples are encouraged.

416 Section on Survey Research Methods Roundtable with Lunch (fee event)

Section on Survey Research Methods
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

Characterization, Evaluation, and Management of Systemic Risk in Surveys

✱John L. Eltinge, Bureau of Labor Statistics, OSMR, PSB 1950, BLS, 2 Massachusetts Avenue NE, Washington, DC 20212, *Eltinge.John@bls.gov*

Key Words: Complex and tightly coupled systems, Constraints, Feedback loops, Operational risk, Paradata, Survey mangement

Survey methodology often emphasizes risk factors that arise at relatively fine levels of aggregation. For many surveys, however, it is important to consider risks that occur at a systemic level (e.g., risks arising from operational constraints; changes in the target population; changes in collection and production systems; loss of personnel; and limitations on timely and accurate survey process information). This roundtable will explore constructive steps survey organizations can take to characterize, evaluate, and manage systemic risks, with emphasis on three questions: 1. What are the most important systemic risks encountered by your survey organization?; 2. What are good approaches you have found to ameliorate these risks?; 3. What are practical ways to communicate these risks—and your program of risk management—to your primary stakeholders?

417 Social Statistics Section Roundtable with Lunch (fee event)

Social Statistics Section
Wednesday, August 6, 12:30 p.m.–1:50 p.m.

From Data to Action: A Case Study To Disseminate and Communicate Children's Health Statistics to Multiple Stakeholders Through an Online Data Tool

✱Kathleen S. O'Connor, Centers for Disease Control and Prevention, National Center for Health Statistics, 3311 Toledo Road, Room 2114, Hyattsville, MD 20782-2003, *koconnor1@cdc.gov*; Christina D. Bethell, Oregon Health & Science University; Stephen J. Blumberg, Centers for Disease Control and Prevention

Key Words: key indicators, data use, data dissemination, children's health, impact of improved data access

Significant investments are made to collect survey data on the health and well-being of children, parents, and families. This roundtable describes an online direct data query tool to improve access to over 60 health statistics. The Data Resource Center (DRC) for Child & Adolescent Health, a project of the Child and Adolescent Health Measurement Initiative (CAHMI) at the Oregon Health & Science University, was designed to enhance data access and use for audiences with various degrees of statistical expertise. The DRC database includes data from several large cross-sectional surveys conducted by the Centers for Disease Control and Prevention. CAHMI works in a collaborative partnership with the Federal Maternal and Child Health Bureau, National Center for Health Statistics, and others to achieve its mission.

418 Spatial Models of Loa loa Disease ●▲

JASA, Applications and Case Studies
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Bivariate Binomial Spatial Modeling Loa loa Prevalence In Tropical Africa

✱Ciprian M. Crainiceanu, Johns Hopkins University, 615 N. Wolfe Street E3636, Department of Biostatistics, Baltimore, MD 21205, *ccrainic@jhsph.edu*; Peter J. Diggle, Lancaster University; Barry Rowlingson, Lancaster University

Key Words: Geostatistics, low-rank, thin-plate splines

We present a state-of-the-art application of smoothing for dependent bivariate binomial spatial data to Loa loa prevalence mapping in West Africa. This application is special because it starts with the non-spatial calibration of survey instruments, continues with the spatial model building and assessment and ends with robust, tested software that will be used by the field scientists of the World Health Organization for online prevalence map updating. From a statistical perspective several important methodological issues were addressed: (a) building spatial models that are complex enough to capture the structure of the data but remain computationally usable; (b) reducing the computational burden in the handling of very large covariate data sets; (c) devising methods for comparing spatial prediction methods for a given exceedance policy threshold.

419 Disease Prediction and Model Validation: Tools for the Real World ●▲

Section on Statistics in Epidemiology, Section on Statisticians in Defense and National Security, WNAR, Section on Health Policy Statistics, Biometrics Section

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Statistical Applications in Elucidating the Epidemic of Obesity and the Metabolic Syndrome in Children and Adolescents

✱Shumei S. Sun, Virginia Commonwealth University, Department of Biostatistics, 730 E Broad Street, Richmond, VA 23298-0032, *ssun@vcu.edu*

Key Words: metabolic syndrome, longitudinal model, ATP III, The Fels Longitudinal Study

Serial data from the Fels Longitudinal Study (FLS) were used to link childhood obesity and other risk factors to the metabolic syndrome later in life to determine the onset of divergence in BMI, waist circumference, and blood pressure among adults with and without the metabolic syndrome and to provide age- and sex-specific childhood values for risk factors that predict the metabolic syndrome. Longitudinal random effects modeling was applied to a discovery sample within the FLS population who met the ATP III criteria for the metabolic syndrome and a sample of randomly selected men and women within the FLS population matched for age and sex who did not meet the criteria for the metabolic syndrome. We then performed a study of predictive accuracy in a validation sample of participants in the FLS using logistic regression and receiver operating characteristic curve analysis.

Evaluation of Prognostic Models Using Area Between Curves (ABC)

✱Mithat Gonen, Memorial Sloan-Kettering Cancer Center, 307 East 63rd street 3rd Floor, New York, NY 10065, *gonenm@mskcc.org*

Key Words: predictive accuracy, discrimination, calibration, censored data, local regression

Evaluation of prognostic models usually involve measuring the discriminatory power using the ROC curve and the AUC as well as an assessment of calibration by comparing predictions with observations. The methods developed in the context of binary data and logistic regression require some adjustments when used in presence of censored data. This talk will discuss the use of ROC curves with censored data and computation of the AUC with Cox models as well as with Kaplan-Meier estimates. Assessment of calibration will be performed by an estimate of the calibration curve using local regression. Area between curves (ABC) will also be discussed as a measure of calibration of prognostic models. Staging of colon cancer using national databases will be used as an example.

Simple Algorithms To Predict Kidney Disease: Ready To Be Used in the Real World?

✱Heejung Bang, Weill Medical College of Cornell University, 411 East 69th St., New York, NY 10021, *heb2013@med.cornell.edu*; Madhu Mazumdar, Weill Cornell Medical College

Key Words: kidney disease, SCORED, risk score, logistic regression, NHANES, validation

Chronic kidney disease is a major public health problem. Given asymptomatic nature, individuals and physicians may be unaware of the condition. Identifying individuals in early stage would be a first step in preventing progression to end stage and reducing morbidity and mortality. Bang et al. (2007, 2008) developed and validated a novel screening model/guideline, SCreening for Occult REnal Disease (SCORED). Logistic regression and split-sample method were used to derive a model and scoring system from the National Health and Nutrition Examination Surveys. Simple screening questionnaire for kidney disease has been shownn to outperform current clinical practice guidelines. Risk and screening models are a great research field in which statisticians can contribute and lead in the real world.

420 Design and Analysis of Industrial Split-Plot Experiments: Best Practices ●

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Aspects in the Analysis of Split-Plot Experiments

✱Oliver Schabenberger, SAS Institute Inc., SAS Campus Dr., Cary, NC 27513, *oliver.schabenberger@sas.com*

Key Words: mixed models, degrees of freedom, variance components, industrial experimentation

Split-plot experiments in the industrial and agricultural settings have much in common such as a hierarchical random effects structure, a relatively small number of whole-plot levels, and an analytic focus on treatment contrasts. The historical connection between field experimentation and mixed model analysis is manifest in contemporary tools for mixed model inference. Whereas this creates substantive overlap with applications in industrial experimentation, it also asks for careful evaluation of the effects of fractionating factorials, factor level coding, and model translation. This talk presents insights into the use of the MIXED procedure in SAS/STAT© for the analysis of industrial split-plot experiments; specifically, the accommodation of fractional factorials, methods of variance component estimation, and degree-of-freedom methods.

Designing Optimal Split-Plot and Split-Split-Plot Experiments

✱Peter Goos, Universiteit Antwerpen, Prinsstraat 13, Antwerpen, International 2000 Belgium, *Peter.Goos@ua.ac.be*; Bradley Jones, SAS Institute Inc.

Key Words: D-optimal design, multi-step experimentation, restricted randomization, design construction algorithm

In industrial experimentation there is growing interest in studies that span more than one processing step. Convenience often dictates restrictions in randomization in passing from one processing step to another. When the study involves two processing steps, this leads to split-plot designs. When the study encompasses three processing steps, this leads to split-split-plot designs. We provide algorithms for computing D-optimal split-plot and split-split-plot designs and provide several illustrative examples. We then apply our methods to construct D-optimal alternatives to a previously run split-split-plot design for cheese production and discuss the problem of designing a gas plasma study.

Split-Plot Designs: What, Why, and How

✱Christopher Nachtsheim, Carlson School of Management, 3-245 CarlSMgmt, 321 19th Ave S, Minneapolis, MN 55455, *nacht001@umn.edu*

"All industrial experiments are split-plot experiments." This provocative remark has been attributed to the famous industrial statistician, Cuthbert Daniel, by Box, Hunter, and Hunter (2005) in their well-known text on design of experiments. The importance of split-plot experiments in industry has been recognized since the first half of the 20th century (Yates, 1936). Despite this, many industrial experiments are fielded as split-plot experiments yet erroneously analyzed as if they were completely randomized designs. This is often the case when hard-to-change factors exist and economic constraints preclude the use of complete randomization. There has been a resurgence of research interest in split-plot design in the last decade or so. Our goal here is to review these recent developments and provide guidelines for the practitioner on their use.

421 Statistical Problems in Wireless Sensor Networks ●

Section on Statisticians in Defense and National Security, Section on Statisticians in Defense and National Security

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Robust Target Detection and Localization in Wireless Sensor Networks

✱George Michailidis, The University of Michigan, *gmichail@umich.edu*

Key Words: Wireless networks, Algorithms

Detecting and localizing a target and estimating its signal are fundamental tasks in wireless sensor networks. Performing them efficiently while conserving energy and communications costs often requires distributed processing. We propose an algorithm for local neighborhood voting and show that performing these operations before global data fusion can significantly improve target detection and reduce communications costs compared to standard global methods. We also show that the same algorithm can improve target localization, where methods based on collecting signal data from a small subset of sensors determined by local operations can achieve performance levels similar to methods based on the full data, but at a fraction of the communications cost.

Energy-Performance Issues for Statistical Inference in Large Random Networks

Animashree Anandkumar, Cornell University; Joseph Yukich, Lehigh University; Ananthram Swami, Cornell University; ✱Lang Tong, Cornell University, 384 Frank H.T. Rhodes Hall, Ithaca, NY 14850, *aa332@cornell.edu*

Key Words: distributed detection, markov random fields, minimum energy routing, error exponent, Stabilizing graph functionals

The issues of energy consumed in data fusion of a multi-hop sensor network and the resulting inference performance at the fusion center are considered. The correlation between the sensor measurements is incorporated via a Markov random field model with a distance-dependent correlation strength. The sensor placement is assumed to be i.i.d. according to a given distribution and its effect on energy and performance is studied. For Poisson and uniformly placed sensors, there is an optimal sensor density that maximizes the error exponent subject to a constraint on the average routing energy. This optimal density crucially depends on the ratio between the measurement variances under the two hypotheses and displays a threshold behavior. Below the threshold, the optimal density tends towards infinity. Above the threshold, it is the minimum feasible value delivering the likelihood ratio.

Active Data Acquisition for Mobile Sensor Networks

✱Nicolas Hengartner, Los Alamos National Laboratory, 2166 Loma Linda, Los Alamos, NM 87544, *nickh@lanl.gov*

Key Words: mobile sensor network, active data acquisition, data marshaling

Large area biological air monitoring of metropolitan areas enables early detection of terrorist biological releases. An important statistical question is how best to deploy the air monitors over the region we want to protect. This talk discusses and compares spatial-temporal sampling designs using either moving or fixed air-sampling units for detecting the release of a biological pathogen into the atmosphere and subsequently for mapping its evolution in space and time. While this talk is motivated by problems in homeland security, the discussion and methodology are more generally applicable to environmental air monitoring.

422 Statistics Education in the Health Sciences During the Clinical and Translational Science Era: Bench to Bedside to Populations ●

Section on Teaching Statistics in the Health Sciences, Section on Statistical Consulting, WNAR, Section on Health Policy Statistics

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Teaching Translational Biostatistics Through Partnered Training: A Dual-Mentoring Approach Using the CTSA

✱Laurel A. Beckett, University of California, Davis, Dept of Public Health Sciences, Division of Biostatistics, MS1C, Davis, CA 95616, *labeckett@ucdavis.edu*

Key Words: mentoring, training, collaboration, translational, clinical, students

To prepare biostatistics students for the real world of medical research, we need to train them in collaboration and translational research. At the same time, we also need to train medical researchers and prepare them to interact with biostatisticians. We have taken advantage of the unique opportunities of the CTSA to develop a dual-mentorship system where biostatistics trainees (graduate students and post-docs) are partnered with mentored clinical research trainees (residents, fellows, and junior faculty), under the dual mentorship of a biostatistician and a clinical or bench researcher. The trainees not only learn from both mentors, but also teach each other and build partnerships that can serve as a model for their future research collaborations. I'll give some examples, discuss principles and challenges, and suggest some outcome metrics for success.

Training Investigators via the CTSA Biostatistics Consulting Core: The Tale of the Two Cities

✱Mariza de Andrade, Mayo Clinic, 200 First Sreet SW, Rochester, MN 55906, *mandrade@mayo.edu*

Key Words: CTSA, consulting, infrastructure, training

The current model of service at the CTSA Service Center statistical help desk has MS- and PhD-level support statisticians staffing the office providing advice and instruction to assist clients in performing their own data analysis and in carrying out many aspects of their research projects. These help-desk services are free to the client. In contrast, statistical support through the Division of Biostatistics is more full-service, fee-based, allows for larger and more complex projects/analyses, and provides the client with a long-term statistical collaborator. Thus, in this presentation I will present the steps needed via CTSA Biostatistics Core to help training investigators from the protocol submission to statistical analysis.

Focusing Biostatistics Training on Bench to Beside Translation

✱Shannon K. McWeeney, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, CR 145, Portland, OR 97232, *mcweeney@ohsu.edu*

Key Words: Translational Research, CTSA, NIH Roadmap

The recent formation of a new national consortium, funded through Clinical and Translational Science Awards (CTSAs) is transforming how clinical and translational research is conducted. This has led to a variety of challenges and opportunities for biostatistics training with the integration of these CTSA centers within the University setting. Several examples will be discussed to highlight how training can keep pace with this dynamic new environment to ensure biostatistics is part of the "translational life cycle."

423 Ideas That Make a Difference: Contributions of F. N. David Award Recipients ●▲

Committee on Women in Statistics

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Nonparametric Linkage Tests Using Affected Pedigree Members: The Work of Alice S. Whittemore

✱Mei-Chiung Shih, Stanford University, HRP Redwood Building, 259 Campus Drive, Stanford, CA 94305, *meichiun@stanford.edu*

Key Words: linkage analysis, allele-sharing methods, lod scores

Nonparametric tests based on allele sharing among pedigree members affected with a given disease are widely used in linkage analysis for gene discovery. In this talk I will review two contributions of Alice S. Whittemore in this area. The first one concerns extending allele sharing methods from affected relative pairs to arbitrary sets of affected pedigree members (Whittemore & Halpern, Biometrics 1994, 50:118–127). The second contribution concerns relating nonparametric allele sharing methods to parametric lod scores tests by a unified likelihood framework (Whittemore, *American Journal of Human Genetics* 1996, 59:704–716). The impact of this work on the field of statistical genetics will also be discussed.

Comparing Multiple Treatments: Importance of the Overall Pattern of Results

✱Juliet P. Shaffer, University of California, Berkeley, Department of Statistics, 367 Evans Hall # 3860, Berkeley, CA 94720-3860, *shaffer@stat.berkeley.edu*

Key Words: Multiple comparisons, Multiple testing, Pairwise comparisons, Familywise error rate, False discovery rate, Newman-Keuls method

Most research studies are concerned with a number of different outcomes. In giving statistical advice, I soon became aware of the importance of considering the total number of inferences and the increase in the probability of drawing false conclusions if methods don't take this multiplicity into account. Most of my research has been in this area. The usual aim is to reject as many false hypotheses as possible while controlling some overall probability of rejecting true hypotheses. Some time ago, based on researcher reactions, I became interested in looking also at the pattern of rejections, so that the total conclusion would be as interpretable as possible. I've taken up this research again recently with some new approaches, and will describe the original motivation, the development of concepts in this area, and some results and possible future developments.

On Collaborations and Contributions: A Personal Perspective

✱Nan Laird, Harvard School of Public Health, *laird@hsph.harvard.edu*

I will discuss the interplay between contributions to statistical methodology and collaborations, drawing on examples such as the EM algorithm, the mixed effects model for longitudinal data and robust tests for genetic association. Collaborations include measuring levels of air pollutants, quantifying the growth of lung function in children, and finding genes for Alzheimer's disease.

A Randomly Reinforced Urn Design

✱Nancy Flournoy, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211, *flournoyn@missouri.edu*

Key Words: response-driven designs, asymptotic theory, branching processes, clinical trials, degenerate information, stochastic processes

I will begin with a brief overview of my involvement in the development of adaptive designs. Then I will describe the Randomized Reinforcement Design. Urn models are useful when randomization is important. Treatment assignments correspond to the color of a drawn ball. If balls are added or taken away depending on outcomes, the urn is "response-driven." The Randomized Reinforcement Urn (RRU) is optimal in the sense that it assigns patients to the best treatment with probability converging to one. No other urn model in the literature has this property. The RRU was first studied by Durham, Flournoy and Li (1998) for binary random variables. However, because allocation proportions degenerate to 0 or 1 (or to a random variable if treatments are equal), theory that permits testing treatment differences has only recently been obtained (May & Flournoy, 2008).

424 Multiscale Methods for Overcomplete Representations ▲

Section on Bayesian Statistical Science

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Time Series Analysis with Overcomplete Representations

✱Guy Nason, University of Bristol, BS8 1TW United Kingdom, *g.p.nason@bristol.ac.uk*; Alessandro Cardinali, University of Bristol

Key Words: time series analysis, locally stationary series, wavelets, wavelet packets, stationarity

The locally stationary wavelet processes (LSW) were introduced in 2000 as an alternative to locally stationary Fourier processes (LSF). This talk first considers testing for second-order stationarity of time series, first by using tests based on overcomplete wavelet and local Fourier analyses and then by a more general analyses based on overcomplete wavelet packet representations. This leads us to the definition of locally stationary wavelet packet processes, a very general model for locally stationary time series. We discuss methods for fitting such processes and the difficulties that arise. We show these methods for analyses and modeling of both stationary and nonstationary series.

Bayesian Rules for Nonparametric Regression Estimation Based on Overcomplete Representations

✱Marianna Pensky, University of Central Florida, 32816, *mpensky@pegasus.cc.ucf.edu*

Key Words: Bayesian inference, Nonparametric regression,

Nonparametric regression estimators based on orthonormal transformations (e.g., Fourier or wavelet) have been studied in statistical literature over the last two decades. Methodologies usually rely on the fact that i.i.d. normal variables are transformed into themselves, and model selection is usually carried out via hard or soft thresholding. In engineering applications, however, overcomplete representations (e.g., Gabor frames, short time Fourier transform) have been far more popular. However, the fact that the noise of the transformed vector does not follow i.i.d. model presents new challenges. Model selection algorithms usually rely on convex optimization techniques, and high dimensionality presents a major challenge. The objective of the present research is to explore Bayesian versions of the known model selection algorithms and practical aspects of their implementation.

Multiscale Time-Frequency Representations and Inference for Forensic Speech and Audio Analysis

✱Patrick J. Wolfe, Harvard University, 33 Oxford St Rm MD-129, Cambridge, MA 02138, *wolfe@stat.harvard.edu*

Key Words: Local stationarity, Time series, Signal processing, Multiscale methods, Time-frequency analysis

The variability inherent in speech waveforms gives rise to powerful temporal and spectral dynamics that evolve across multiple scales. As a result, optimal analysis/synthesis schemes may differ as a function of phoneme, prosody, and acoustic environment. However, most current schemes (and therefore algorithms that rely on them) operate at a fixed resolution chosen a priori. In this talk we describe methods for adaptive short-time Fourier analysis that proceed according to measures of time-frequency spread and the statistical notion of local stationarity. Subsequent synthesis is made possible through perfect-reconstruction filter bank methods and a simple modification of overlap-add procedures.

425 The Emergence of Social Data Analysis and Its Impact on the Field of Statistics ●▲

Section on Statistical Graphics, Social Statistics Section, *CHANCE*

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Many Eyes: A Site for Social Data Analysis and Visualization

Fernanda Viegas, IBM; ✱Martin Wattenberg, IBM Research, Visual Communication Lab, 1 Rogers Street, Cambridge, MA 02142, *mwatten@us.ibm.com*

Key Words: visualization

“Many Eyes” is a web site that lets users collect data, create visualizations, and share questions and discoveries. By making this technology available for free to anyone on the internet, we seek to “democratize” visualization and to foster a new, open style of data analysis and statistical sensemaking. The site counts thousands of registered users who have contributed data sets on subjects ranging from bioinformatics, to the Bible, to the 2008 presidential campaign. At its core, Many Eyes is a bet on the power of visualization to spark statistical analysis–a bet that has paid off in a surprising number of ways, ranging from artistic creation to citizen activism and journalism.

StatCrunch: From Statistical Software to a Portal for Social Data Analysis

✱Webster West, Integrated Analytics LLC, 17302 Sioux Springs Drive, Chief Executive Officer, College Station, TX 77845, *support@statcrunch.com*

Key Words: online, statistical software, social data analysis

For many years, StatCrunch has been available as an online statistical software package, and the software has developed a large user base in the arena of statistical education. Recently, the web site surrounding the StatCrunch application has underwent a tremendous overhaul. The site now allows users to actively share their data, results and reports with the rest of the world, and it provides a forum to discuss these items with other interested parties. In short, StatCrunch users can now create rich searchable content for social data analysis. In this talk, the potential positive impacts of this approach to software delivery on the field of statistics will be discussed. A specific focus will be given to the field of statistical education.

426 Best of the *Annals of Applied Statistics* ●

IMS, WNAR

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Stochastic Modeling in Nanoscale Biophysics: Subdiffusion Within Proteins

✱Samuel Kou, Harvard University, Department of Statistics, Cambridge, MA 02138, *kou@stat.harvard.edu*

Key Words: single-molecule experiments, protein conformational dynamics, memory kernel, generalized Langevin equation, autocorrelation function, protein conformational dynamics, memory kernel, generalized Langevin equation, autocorrelation function

Advances in nanotechnology have allowed scientists to study biological processes on an unprecedented nanoscale molecule-by-molecule basis, opening the door to addressing many important biological problems. A phenomenon observed in recent nanoscale single-molecule biophysics experiments is subdiffusion, which largely departs from the classical Brownian diffusion theory. In this paper, by incorporating fractional Gaussian noise into the generalized Langevin equation, we formulate a model to describe subdiffusion. In addition to its analytical tractability and clear physical underpinning, the model is capable of explaining data collected in fluorescence studies on single protein molecules. Excellent agreement between the model prediction and the single-molecule experimental data is seen.

Of Mice and Men: Sparse Statistical Modeling in Genomics

✱Mike West, Duke University, Department of Statistical Science, Durham, NC 27708-0251, *mike@stat.duke.edu*

I discuss applied statistical modeling of gene expression data to evaluate experimentally defined biomarkers in human observational data sets. The applied context—cross-study, cross-species projection of expression-based biomarkers—is central to emerging interests in clinical applications of genomics. The AOAS example in cardiovascular genomics involves a gene-environment interaction experiment that links gene expression to disease states and known risk factors. Expression signatures that proxy such risk factors are then evaluated in a coupled exploratory analysis of human data: how do these mouse-model signatures translate to humans? The study highlights a range of applied statistical and genomic data analysis issues, including model specification, computational questions, and model-based correction of experimental artifacts in DNA microarray data.

Should the Democrats Move to the Left on Economic Policy?

✱Andrew Gelman, Columbia University, 1016 SSW, Statistics Department, 1255 Amsterdam Avenue, New York, NY 10027, *gelman@stat.columbia.edu*; Cexun J. Cai, Columbia University

Key Words: median voter, presidential election, public opinion, spatial model, voting

Could John Kerry have gained votes in the recent Presidential election by more clearly distinguishing himself from George Bush on economic policy? At first thought, the logic of political preferences would suggest not. In a multidimensional setting, however, or when voters vary in their perceptions of the parties' positions, a party can benefit from putting some daylight between itself and the other party on an issue where it has a public-opinion advantage (such as economic policy for the Democrats). We set up a model based on survey data on voters' perceptions of their own positions and those of the candidates in 2004. Under our model, it turns out to be optimal for the Democrats to move slightly to the right but staying clearly to the left of the Republicans' current position on economic issues.

427 Biometrics Editors Invited Session

ENAR, WNAR, Biometrics Section

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Variable Selection for Model-Based, High-Dimensional Clustering and Its Application to Microarray Data

Sijian Wang, The University of Michigan; ✱Ji Zhu, The University of Michigan, Dept of Statistics, University of Michigan, 1085 South University, Ann Arbor, MI 48109-1107, *jizhu@umich.edu*

Key Words: EM Algorithm, High Dimension Low Sample Size, Microarray, Model-Based Clustering, Regularization, Variable Selection

Variable selection in high-dimensional clustering analysis is an important yet challenging problem. In this paper, we propose two methods that simultaneously separate data points into similar clusters and select informative variables that contribute to the clustering. Our methods are in the framework of penalized model-based clustering. Unlike the classical L1-norm penalization, the penalty terms that we propose make use of the fact that parameters belonging to one variable should be treated as a natural "group." Numerical results indicate that the two new methods tend to remove noninformative variables more effectively and provide better clustering results than the L1-norm approach.

Screening for Partial Conjunction Hypotheses

✱Yoav Benjamini, Tel Aviv University, Department of Statistics and O.R., Tel Aviv, 69978 Israel, *ybenja@tau.ac.il*; Ruth Heller, University of Pennsylvania

Key Words: multiple testing, fMRI, false discovery rate, Microarray analysis, mult- subject study

We address the problem of testing for partial conjunction of hypotheses, that argues that at least u out of n tested hypotheses are false. It offers an in-between approach to the testing of the conjunction of null hypotheses against the alternative that at least one is not, and the testing of the disjunction of null hypotheses against the alternative that all hypotheses are not null. We then address the problem of testing many partial conjunction hypotheses simultaneously using the false discovery rate (FDR) approach. Moreover, we can screen at all levels of conjunctions simultaneously in order to display the findings on a superimposed map and still control an appropriate FDR measure. Needs for such methods have been identified when combining microarray experiments, as well as in functional Magnetic Resonance Imaging (fMRI).

A Flexible and Powerful Bayesian Hierarchical Model for ChIP-Chip Experiments

✱Raphael Gottardo, University of British Columbia, 333-6356 Agricultural Road, Vancouver, BC V6T1Z2 Canada, *raph@stat.ubc.ca*; Wei Li, Harvard University; Evan Johnson, Brigham Young University; Shirley Liu, Harvard University

Key Words: Bayesian hierarchical model, Empirical Bayes, Markov chain Monte Carlo, Multiple testing, Tiling Arrays, Spatial Statistics

Chromatin-immunoprecipitation microarrays (ChIP-chip) that enable researchers to identify regions of a given genome that are bound by specific DNA-binding proteins present new challenges for statistical analysis. We propose a method called BAC (Bayesian analysis of ChIP-chip) to detect transcription factor bound regions, which incorporate the dependence between probes while making little assumptions about the bound regions (e.g., length). BAC is robust to probe outliers with an exchangeable prior for the variances, which allows different variances for the probes but still shrink extreme empirical variances. Bound regions are detected using posterior probabilities computed from the joint posterior distribution of neighboring probes. We show that these posterior probabilities are well calibrated and can be used to obtain an estimate of the false discovery rate.

428 Bioinformatics ●

WNAR, Biopharmaceutical Section, Biometrics Section

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Survival Analysis of Microarray Gene Expression Data

✱Jianguo (Tony) Sun, University of Missouri, Columbia, MO , *sunj@missouri.edu*; Qiang Zhao, Texas State University; Ping Yao, University of Missouri-Columbia

Key Words: Gene expression data, censoring, dimension reduction, survival analysis

Statistical analysis of microarray gene expression data has recently attracted a great deal of attention. One problem of interest is to relate genes to survival outcomes of patients with the purpose of building regression models for the prediction of future patients' survival based on their gene expression data. This talk will discuss issues that occur in fitting these data to the commonly used survival models including the proportional hazards or Cox model. Several models and their corresponding inference approaches will be considered.

Reconstructing Gene Regulatory Networks from Gene Expression Data and Biological Prior Knowledge

✱Dirk Husmeier, Biomathematics & Statistics Scotland, JCMB, Room 3606, The King's Buildings, Edinburgh, International EH9 3JZ United Kingdom, *dirk@bioss.ac.uk*

Key Words: Gene regulatory networks, Bayesian networks, KEGG, microarrays, MCMC, flow cytometry

The objective of the talk is the discussion of a method for improving the reconstruction of gene regulatory networks from microarray data by the systematic integration of biological prior knowledge. Our approach is based on the Bayesian paradigm whereby a prior distribution over network structures is derived from partial knowledge of signaling pathways, as obtained from databases such as KEGG. The hyperparameters of this distribution represent the weights associated with the prior knowledge relative to the data. We have derived and tested an MCMC scheme for sampling networks and hyperparameters simultaneously from the posterior distribution, thereby automatically learning how to trade off information from the prior and the data. We have assessed the viability of the proposed method on cytometry data obtained from the RAF signaling pathway.

Modeling Splice-Junction Microarrays

✱Kasper D. Hansen, University of California, Berkeley, Division of Biostatistics, School of Public Health, 101 Haviland Hall, #7358, Berkeley, CA 94720-7358, *khansen@stat.berkeley.edu*

Key Words: microarrays, non-linear regression

Alternative splicing is hypothesized to play a major role in gene diversity in higher organisms. Splice-junction microarrays are microarrays designed to monitor alternative isoforms as opposed to the gene-level measurements common in ordinary gene expression microarrays. Tto deconvolve the microarray measurements, we are proposing a nonlinear interpretable model with parameters corresponding to the isoform fold changes and relative abundances. The model is fit using nonlinear regression. We apply the model to study nonsense mediated degradation in Drosophila melanogaster.

429 Bayesian Modeling of Extreme Events

Section on Risk Analysis, Section on Statistics and the Environment, Section on Statistics in Epidemiology

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Semiparametric Functional Estimation Using Quantile-Based Prior Elicitation

✱Elijah Gaioni, University of Connecticut, *elijah.gaioni@uconn.edu*; Dipak Dey, University of Connecticut; Mircea Grigoriu, Cornell University

Key Words: semiparametric, functional, prior, elicitation

A methodology by which inconsistent prior information can be used to perform functional estimation is presented. Sharp qualitative information consisting of the functional form of the likelihood is assumed known. It is also assumed that vague quantitative information in the form of multiple possible quantiles is available. An optimality criterion is then applied for the purpose of determining a predictive distribution consistent with the expert provided information. The prior distribution is estimated semiparametrically, where an adaptive method for selecting basis elements is used to limit the computational difficulties associated with the solution of this problem.

Bayesian Hierarchical Modeling for Extreme Values Observed Over Space and Time

✱Huiyan Sang, Duke University, 910 Constitution Dr, Apt 708, Durham, NC 27705, *hs37@duke.edu*; Alan Gelfand, Duke University

Key Words: Multivariate extreme values, Spatial-temporal model, Climate risk analysis

In this talk, I will present our hierarchical modeling approach for explaining a collection of spatially referenced time series of extreme values. The univariate distributions of extreme values are extended to higher dimensions using latent multivariate Markov random field models specified through coregionalization, which allows the interpretation of high dimensional extreme value analysis including the nature of spatial association and the nature of temporal trend. By relaxing the assumption of conditional independence in the hierarchical models, we extend our approach to describe extreme values with a smoothed spatial process, which can be used in spatial interpolation with extremes.

Bayesian Model Selection of the Farlie-Gumbel-Morgenstern Copula for Describing Two Generalized Extreme Value Variables

✱Vered Madar, SAMSI, 19 T.W. Alexander Drive, P.O. Box 14006, Research Triangle Park, NC 27709-4006, *madar@samsi.info*

Key Words: Bayesian model selection, Copula, Farlie-Gumbel-Morgenstern, generalized extreme value variables

The Bivariate Farlie-Gumbel-Morgenstern Copula $F(X)G(X) + \alpha (1-F(X)) (1-G(X))$ (Johnson and Kotz 1975) has a bad reputation for providing restricted range of dependence (Joe 1997). Yet the simple linear structure of this copula is very appealing, and with slight changes one can increase further its range of dependence (Guven and Kotz 2008). We suggest a Bayesian model selection procedure for this copula when describing the joint distribution of two generalized extreme value variables.

A Bayesian Framework for Adversarial Risk Analysis

✱Jesus Rios, SAMSI, *jrios@samsi.info*; David Rios, Universidad Rey Juan Carlos; David Banks, Duke University

Key Words: Adversarial risk analysis, Game theory, Asymmetric prescriptive/descriptive decision analysis, Bayesian analysis

Applications in counterterrorism and corporate competition have led to the development of new methods for the analysis of decisionmaking when there are intelligent opponents and uncertain outcomes. This field represents a combination of statistical risk analysis and classical game theory, and is sometimes called adversarial risk analysis. We describe several approaches to adversarial risk problems, providing a unified framework for analysis aimed at prescribing advice to one of the participants. The key issue in our framework is the assessment of the probabilities of adversaries' actions. We assume that adversaries are expected utility maximizers and, therefore, uncertainty in their actions stem from our uncertainty about their utilities and probabilities when used to analyzed the adversaries' decision problems.

Thresholding for Multivariate Extreme Values

✱Kobi A. Abayomi, Duke University, 223A Old Chemistry, Dept of Statistics, Durham, NC 27707, *kaa21@duke.edu*

Key Words: thresholding, extreme value, environmental statistics, model distance, kullback leibler, environmental hazards, disasters

Extreme Value Distributions (EVDs) are often used, in environmental data settings, to characterize processes that govern natural disasters. Multivariate EVDs can be applied to vector data, which may cover a suite of natural disasters. Thresholding in the univariate case is selecting the scalar cutoff where the extreme value theory applies; in the multivariate case the threshold is a multidimensional vector, or contour. I suggest multivariate thresholding via maximization of tail dependence, by taking a model based minimization of the distance between empirical and parametric dependency.

430 Enhancing Development of Genomewide Association Methods ●▲

Biometrics Section, Biopharmaceutical Section, Section on Statistics in Epidemiology, WNAR
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

On Combining Family-Based and Population-Based Case-Control Study in Assessing Genetic Association with Disease Risk

✱Yingye Zheng, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N, M2-B500, Seattle, WA 98033, *yzheng@fhcrc.org*; Patrick Heagerty, University of Washington; Li Hsu, Fred Hutchinson Cancer Research Center

Key Words: Family study, case-control study, genetic association study, conditional likelihood

Pooling data collected from difference sources is a cost-effective and time-efficient approach for enhancing the power to detect weak-to-modest genetic effects and gene*gene and gene*environment interactions. However the analysis is complicated when data are collected under different study designs such as family-based and population based case-control study. We developed a conditional marginalized likelihood based procedure to combine data from both population and family based sources and account for different ascertainment schemes. The work is motivated by the ongoing GWA in the colorectal cancer family registry.

Multiple Comparisons Overestimation Bias in Association Studies

✱Neal Jeffries, National Heart, Lung, and Blood Institute, Office of Biostatistics Research, Bethesda, MD 20912, *neal.jeffries@nih.gov*

Key Words: multiple comparisons, bias, bootstrap, effect size

It is widely appreciated that genomewide association studies often yield overestimates of the association of a marker with disease when attention focuses upon the marker showing the strongest relationship. The most common reason given for this observation is that the examination of the most extreme test statistic is often conditional upon first observing a significant p-value for the marker. A second, less appreciated reason is described herein; under common circumstances it is the multiple testing of markers that directly leads to bias in the estimated effect size associated with the most extreme test statistic. This multiple comparisons bias is separate from that arising from conditioning on a significant p-value and may often be a more important factor. An analytic description of the bias, identification of exacerbating factors, and a bootstrap approach to bias reduction are given.

Multivariate Methods for Detecting and Displaying Population Structure in GWAS

✱Kelci J. Miclaus, North Carolina State University/SAS Institute, 1725 TW Alexander Dr, #802, Durham, NC 27703, *kelci.miclaus@jmp.com*; Russ Wolfinger, SAS Institute Inc.; Wendy Czika, SAS Institute Inc.

Key Words: Population Structure, MDS, PCA, Genetic Distance, GWAS

Understanding the complex structural aspects of genetic data is critical for the analysis of Genome-Wide Association Studies. Exploration into the use of common genetic distance measures combined with multivariate techniques such as Multidimensional Scaling (MDS) and Principal Components Analysis (PCA) provides insight for determining the presence of population structure that may result in spurious associations. The use of Single Nucleotide Polymorphisms (SNPs) that exhibit Hardy-Weinberg Disequilibrium as substructure informative markers is examined in theory and with empirical study in this work. We present an in depth analysis of MDS and PCA for visualizing population structure; along with the behavior of these methods in conjunction with SNP selection choices. Data from the HapMap Project along with simulated data will be utilized.

Multi-SNP Association Mapping Using Bayesian Regression and Shrinkage Priors

✱Yongtao Guan, The University of Chicago, 920 E. 58th Street - CLSC 412, Chicago, IL 60637, *ytguan@uchicago.edu*; Matthew Stephens, The University of Chicago

Key Words: Bayesian regression, Shrinkage prior, Multi-SNP, Association, MCMC

Whole genome association studies are currently genotyping thousands of individuals at hundreds of thousands of markers, in an effort to identify variants affecting phenotypes related to human health. Conventional analyses test each marker, one at a time, for association with phenotype. Here we describe a fully Bayesian approach that analyses all markers jointly, and uses a novel prior to avoid over-fitting. Despite the size of the problem, a simple MCMC scheme gives helpful results in only a few hours, and provide improved inference compared with standard single-SNP analyses. Methods described here are implemented in a software package, BIMBAM, available from Stephens Lab web site.

A Robust Test for Two-Stage Design in Genome-Wide Association Studies

✱Minjung Kwak, National Heart, Lung, and Blood Institute, *kwakm2@nhlbi.nih.gov*; Jungnam Joo, National Heart, Lung, and Blood Institute; Gang Zheng, National Heart, Lung, and Blood Institute

Key Words: genome-wide association study, genetic model selection, robust

Genetic models for single nucleotide polymorphisms with true associations are usually unknown in genome-wide association studies. In this case, existing methods (allele-based test or trend test) for analysis of the two-stage design are not robust across the four genetic models. The robust test, MAX (the maximum of the trend tests optimal for various genetic models), is statistically difficult to apply for the two-stage design with the joint analysis. We study a simple robust procedure with genetic model selection. Our results show that, when the optimal trend test has about 80% power for a known genetic model, existing methods for analysis of the two-stage design have power ranging from 20% to 80% across the four genetic models, while our robust procedure has power consistently from 69% to 80% across the four genetic models.

431 DMC Members: What Do They Hang Their Hats On? ●▲

Biopharmaceutical Section, Biometrics Section
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Beyond Standard Reports

✱Vipin Arora, Takeda Pharmaceuticals North America, One Deerfield Parkway, Deerfield, IL 60015, *vipin.arora@gmail.com*

Key Words: DMC, Independent Statistician, creative

Data Monitoring Committee (DMC) has responsibility to make a recommendation based on the submitted data. DMC relies heavily on the data package submitted. However, DMC is often looks beyond the submitted data package to assess alternate scenarios that may lead to a different decision/recommendation. The resources used by DMC may include Independent Statistician, Literature available on a similar class of compounds or therapies. Tools like simulation are also used to look beyond the observed data and predict the safety and efficacy of the investigative compound. Such indulgence by DMC is extremely critical and needs resources. The security walls between Independent Statistician and clinical trial team limits the data sharing and poses an operational problem. To protect the integrity of the trial, DMC needs to be creative and proactively think of using alternative resources.

Effective Interactions of the DMC, Steering Committee, and Sponsor in a Long-Term Prevention Trial

✱Michelle McNabb, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285, *ma.mcnabb@lilly.com*; Janet Wittes, Statistics Collaborative, Inc.

Key Words: DMC, clinical trials, data monitoring

Long-term prevention trials present special data monitoring challenges and opportunities. This is especially true when medical advancements in disease assessment and treatment are rapidly evolving. The RUTH (Raloxifene Use for the Heart) trial studied over 10,000 post-menopausal women at high risk for cardiovascular events for seven years. Effective communication between the DMC, steering committee and sponsor was essential to preserving the scientific integrity of the trial while allowing for modifications to the trial in response to new information external to the trial. Recommended practices based upon the RUTH trial will be presented from the trial statistician's and DMC chair's perspectives.

Supporting the DMC To Fulfill Their Responsibilities Through Clear, Comprehensive, and Carefully Constructed Reports on Accumulating Interim Data

✱Marian R. Fisher, University of Wisconsin-Madison, 610 Walnut Street, WARF Room 215, Sch of Med & Pub Hlth, Dept of Biostat & Med Info, Madison, WI 53726, *fisher@biostat.wisc.edu*

Key Words: data monitoring committee, independent statistical data analysis center, interim monitoring, DMC reports, incomplete event adjudication, blinded sample size change

Interim data inherently require thought in terms of integration of various sources for timeliness, availability, cleanliness, and consistency. Making a decision to have a DMC evaluate accumulating interim data for safety and efficacy requires preparation of reports based on interim data. Members of the DMC include clinicians and others with collaborative research expertise so DMC reports must be constructed to present analysis results effectively and efficiently. Examples of analysis of time-to-event data with incomplete event adjudication and of a change in sample size in a blinded fashion that allowed flexibility and preserved the overall type I error rate will be presented.

Depending on the Independent Statistician ...

✱Jennifer Schumi, Statistics Collaborative, Inc., 1625 Massachusetts Avenue NW, Suite 600, Washington, DC 20036, *jennifer@statcollab.com*; Matt Downs, Statistics Collaborative, Inc.

Key Words: Clinical Trials, Data Monitoring Committees, Interim Analysis, Unblinded data, Safety Monitoring

FDA's Guidance on DMCs suggests an external or "independent" statistician report unblinded data. Communication between the sponsor, DMC, and reporting statistician can take many forms, some more "independent" than others. In many cases, trials require neither a sophisticated DMC nor an experienced reporting group and the simplest models suffice. Sometimes, however, something happens that requires agility and insight from the reporting statistician, and careful reasoning from the DMC - unanticipated serious adverse events, related data from external trials, or subgroup effects. Drawing from our experience reporting to DMCs for pharmaceutical and biotech sponsors, we will discuss our philosophy of independence, which requires aggressive curiosity and thoughtful reporting from the external statistician, open exchange of information from the sponsor, and measured judgment from the DMC.

432 Student Paper Competition: Bayesian Computational Statistics

Section on Bayesian Statistical Science
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Stochastic Approximation and Newton's Estimate of a Mixing Distribution

✱Ryan Martin, Purdue University, Department of Statistics, 250 North University Street, West Lafayette, IN 47907, *martinrg@stat.purdue.edu*; Jayanta K. Ghosh, Purdue University and Indian Statistical Institute

Key Words: mixture models, empirical Bayes, stochastic approximation

Recent scientific advances have increased the need for efficient methods for estimating high-dimensional mixing distributions. One major area of application is to microarray data on gene expression. In this problem, the parameters represent the effects of the individual genes and an empirical Bayes analysis proceeds by estimating the prior, or mixing distribution, based on data observed from the mixture. Newton (2002) proposed a very fast recursive algorithm for nonparametric estimation of the mixing distribution. Simulations show that this estimate performs well compared to the NPML and NP Bayes estimates, based on accuracy and computational efficiency. For finite mixtures, the algorithm is a special case of stochastic approximation (SA) and its asymptotic properties are derived from standard SA arguments, involving Lyapunov functions and ODE stability theory.

Default Priors and Efficient Posterior Computation in Bayesian Factor Analysis

✱Joyee Ghosh, Duke University, 943 Morreene Road, Apt. B 16, Durham, NC 27705, *joyee@stat.duke.edu*; David Dunson, National Institute of Environmental Health Science

Key Words: Bayes factor, Covariance structure, Latent variables, Parameter expansion, Selection of factors, Slow mixing

Factor analytic models are widely used in social sciences and are also useful for sparse modeling of the covariance structure in multivariate data. Normal priors for factor loadings and inverse gamma priors for residual variances are a popular choice because of their conditionally conjugate form. However, such priors require elicitation of many hyperparameters and tend to result in poorly behaved Gibbs samplers. In addition, one must choose an informative specification, as high variance priors face problems due to impropriety of the posterior. This article proposes a default, heavy tailed prior specification, which is induced through parameter expansion while facilitating efficient posterior computation. We also develop an approach to allow uncertainty in the number of factors. The methods are illustrated through simulated examples and epidemiology and toxicology applications.

Objective Bayesian Model Selection in Gaussian Graphical Models

✱James G. Scott, Duke University, Duke University, Box 90251, Durham, NC 27708, *james@stat.duke.edu*; Carlos M. Carvalho, The University of Chicago

Key Words: Gaussian graphical models, covariance selection, Bayesian model selection, multiple testing, Stochastic search

This paper presents a default model-selection procedure for Gaussian graphical models that involves two new developments. First, we develop an objective version of the hyper-inverse Wishart prior for restricted covariance matrices, called the HIW g-prior, and show how it corresponds to the implied fractional prior for covariance selection using fractional Bayes factors. Second, we apply a class of priors that automatically handles the problem of multiple hypothesis testing implied by covariance selection. Numerical experiments show that these priors strongly control the number of false edges included in the model, thereby automatically rewarding sparsity. We demonstrate our methods on a variety of simulated examples, concluding with a real example analyzing covariation in mutual-fund returns that shows our objective-Bayes methods can yield superior predictive performance.

Density Estimation for Bivariate Angular Data Using a von Mises Distribution and Bayesian Nonparametrics

✱Kristin P. Lennox, Texas A&M University, Department of Statistics, 3143 TAMU, College Station, TX 77843-3143, *lennox@stat.tamu.edu*; David B. Dahl, Texas A&M University; Marina Vannucci, Rice University; Jerry Tsai, Texas A&M University

Key Words: Circular data, Dihedral angles, Dirichlet process mixture model, Protein conformational angles, Torsion angles

Interest in modeling protein backbone conformational angles has prompted the development of bivariate angular distributions, as well as frequentist methods for fitting mixtures of these distributions to data. We present a Bayesian approach to density estimation for bivariate angular data that uses a Dirichlet process mixture (DPM) model and the bivariate von Mises distribution. We derive the necessary full conditional distributions to fit such a model, as well as the details for sampling from the posterior predictive distribution. This method is used to estimate the main-chain torsion angle distributions from the immunoglobulin protein structure family.

433 U.S. Census Bureau: Census 2010 ●

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

An Overview of the 2010 Redesign Program at the U.S. Census Bureau

✱Ruth A. Killion, U.S. Census Bureau, *ruth.ann.killion@census.gov*; James Farber, U.S. Census Bureau; Padraic Murphy, U.S. Census Bureau

Key Words: Demographic Survey, Sample Design, Master Address File

Following the decennial census, the Census Bureau and other agencies that sponsor major demographic surveys, such as the Current Population Survey, perform a redesign of those surveys. The overarching goals of the redesign are to ensure that survey requirements are met, to incorporate new or changing requirements, and to assess and improve statistical methods and operational processes in the ongoing survey programs. These goals are met through a comprehensive and integrated program of evaluation, research, development, and implementation. This paper will give an overview of the 2010 Redesign Program at the Census Bureau. We will highlight research on potential changes in statistical methods, such as replacing the census long-form data with the American Community Survey, using a single frame for sample selection, and incorporating administrative records in the sample design.

An Overview of Primary Sampling Units (PSUs) in Multistage Samples for Demographic Surveys

✱Padraic Murphy, U.S. Census Bureau, *padraic.a.murphy@census.gov*

Key Words: sample design, stratified sampling, history of surveys

This paper will present an overview of historical and current issues in the definition, stratification, and selection of Primary Sampling Units (PSUs) in large demographic surveys. By PSUs, we mean the large clusters (usually geographic areas such as counties) that are the sampling frame in the first stage of a multi-stage sample design. Although the concept of a PSU is simple, the details of defining, stratifying, and selecting PSUs can prove to be surprisingly complex. We look at developments pertaining to PSUs over the past half-century, and describe some current problems. The issues discussed include (1) constraints on PSU size and boundaries, (2) choosing "building blocks" for PSUs, (3) methodology for and objectives of PSU stratification, (4) coordination among multiple surveys, and (5) coordination with preceding designs (i.e., maximizing or minimizing overlap.)

Use a MAF-Based Frame for Demographic Household Surveys

✱Xijian Liu, U.S. Census Bureau, *xijian.liu@census.gov*

Key Words: Master Address File, survey frame, coverage

One of the goals of the 2010 Demographic Survey Redesign is to use the continually updated Master Address File as the primary source to develop the sampling frames for current demographic household surveys. To support this goal, the Census Bureau is conducting several evaluations to compare the coverage of a MAF-based frame and the current four frames. The focus is on the overall quality at the national level as well as in two sub-universes: the sub-universe of new addresses added since the last census and the sub-universe currently covered by an area frame that is primarily in rural areas. This paper will present the most recent findings from these evaluations.

Innovations in Survey Redesign at the U.S. Census Bureau

✱James Farber, U.S. Census Bureau, *james.farber@census.gov*; Padraic Murphy, U.S. Census Bureau

Key Words: sample design, data processing, interviewing, cost

The Redesign Program at the Census Bureau ensures that survey requirements are met and new requirements are incorporated for demographic surveys. The redesign also provides an opportunity to evaluate and improve statistical methods and operational processes. However, timing limits the amount of research and innovation that can be done within a single redesign. The new samples must be designed, selected, and fielded for interviewing under a strict schedule. To increase the focus on long-term strategy, the Census Bureau created a permanent Redesign Program. This program is responsible not only for managing the current redesign but also for researching topics that otherwise might not be examined due to tight deadlines in the current redesign. This paper discusses some of these long-term research ideas, and also reviews some innovations of past redesigns and the ensuing benefits.

434 Bootstrap Methods for Complex Problems ▲

Section on Nonparametric Statistics, IMS, Section on Statistical Computing, SSC

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

On Block Bootstrap for Spatial Linear Regression of Irregularly Spaced Data

✱Jun Zhu, University of Wisconsin-Madison, Department of Statistics, 1300 University Avenue, Madison, WI 53706, *jzhu@stat.wisc.edu*

Key Words: Spatial sampling design, Random field, Resampling method

I consider block bootstrap for a spatial linear regression model where the error term is modeled as a random field. To accommodate irregularly spaced sampling sites, the spatial sampling design is assumed to be stochastic and possibly nonuniform. A spatial block bootstrap method is proposed to approximate the distribution of ordinary and penalized least squares estimators of the regression coefficients. Its theoretical and empirical properties are investigated, along with illustrative numerical examples.

Asymptotic Properties of Sample Quantiles from a Finite Population

✱Arindam Chatterjee, Texas A&M University, Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143, *cha@stat.tamu.edu*

Key Words: Finite population, Sampling, Quantiles, Bootstrap, Variance estimation

In this paper we consider the problem of estimating quantiles of a finite population of size N on the basis of a finite sample of size n selected without replacement. We prove the asymptotic normality of the sample quantile and show that the scaled variance of the sample quantile converges to the asymptotic variance under a slight moment condition. We also consider the performance of the bootstrap in this case, proposed by Gross (1980). We show that Gross's bootstrap method fails to be consistent, but a suitably modified version of the bootstrapped quantile converges to the same asymptotic distribution as the sample quantile. Consistency of the modified bootstrap variance estimate is also proved under the same moment conditions.

Variance Estimation for Sample Quantiles Using Smoothed Moving Block Bootstrap

✱Shuxia Sun, Wright State University, 159 MM BLG, Wright State University, 3640 Col. Glenn Hwy, Dayton, OH 45435, *shuxia.sun@wright.edu*

Key Words: moving block bootstrap, sample quantile, strongly mixng, smoothed bootstrap estimator

Under dependence, the asymptotic variance of the normalized sample quantile not only involves the density of the random variable evaluated at the unknown population quantile, but at the same time, an infinite series of lag-covariances of a transformed sequence. In this paper, we compare the performance of three variance estimators, the nonparametric kernel-smoothed estimator, the moving block bootstrap estimator, and the smoothed bootstrap estimator. We show that, as in the independently identically distributed case, the smoothed MBB variance estimator outperforms the other two estimators. A simulation study is presented to compare the empirical performance of these three estimators.

Pooling Block Bootstrap Estimators

✱Arnab Maity, Texas A&M University, Department of Statistics, 3143 TAMU, College Station, TX 77843, *amaity@stat.tamu.edu*; Soumendra N. Lahiri, Texas A&M University

Key Words: Block bootstrap, Resampling, Time series, Variance estimation

The choice of appropriate block length is a very important issue in construction of block bootstrap estimators. Typically, one chooses a block size that minimizes the mean squared error (MSE) of the estimator. However, in many cases the estimator can be very sensitive of the choice block size and may change dramatically even for a small change in block length. In this paper, we construct pooled bootstrap estimators by combining estimates for several block sizes and investigate asymptotic properties of such estimators. It is shown that for a large class of pooled bootstrap estimators, the MSE optimal rate can be attained, without the need to preselect an optimal block size. Results from a simulation study will also be presented to illustrate finite sample properties of the proposed method.

A Tapered Block Bootstrap for Variance Estimation in Time Series Regression

✱Dan Nordman, Iowa State University, Department of Statistics, Iowa State University, Ames, IA 50011, *dnordman@iastate.edu*

Key Words: asymptotic expansion, least squares, resampling

Paparoditis and Politis (2001) developed the tapered block bootstrap (TBB) for estimating the variance of a time series sample mean. This talk considers the TBB for estimating the variance of least squares estimators in more general time series framework with nonstochastic regressors. The original TBB formulation may break down and produce inconsistent variance estimators when regressors are nonconstant (i.e., outside of the sample mean case). An alternative TBB is proposed with a modified resampling mechanism. This TBB version matches the original one in the sample mean case and will produce consistent variance estimators under any nonstochastic regressors. This TBB also carries the desirable bias properties of the original TBB formulation into the regression setting.

435 Least Angle Regression

Section on Statistical Computing, IMS

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Hierarchically Penalized Cox Regression for High-Dimensional Censored Data with Grouped Variables and Its Oracle Property

✱Sijian Wang, The University of Michigan, Dept of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109-2029, *sijwang@umich.edu*; Bin Nan, The University of Michigan; Nengfeng Zhou, The University of Michigan; Ji Zhu, The University of Michigan

Key Words: Cox model, Group variable selection, Lasso, Microarray, Oracle property, Regularization

In many biological and other scientific applications, prediction variables are naturally grouped. For example, in biological applications, assayed genes or proteins are grouped by biological roles or biological pathways. When studying the dependence of survival outcome on these grouped prediction variables, it is desirable to select variables at both the group level and the variable-specific level. In this article, we develop a new method to handle the group variable selection in the Cox proportional hazards model. Our method not only effectively removes unimportant groups, but also maintains the flexibility of selecting variables within the identified groups. We also show that the new method offers the potential for achieving the asymptotic oracle property as in Fan and Li (2001, 2002).

Variable Selection with the Strong Heredity/Marginality Constraint and Its Oracle Property

✱Nam Hee Choi, The University of Michigan, MI 48105, *nami@umich.edu*; Ji Zhu, The University of Michigan

Key Words: Heredity, LASSO, Marginality, Oracle property, Regularization, Variable selection

We extend the LASSO method (Tibshirani, 1996) for simultaneously fitting a regression model and identifying important interaction terms. Unlike most of the existing variable selection methods, our method automatically enforces the heredity/marginality constraint, i.e., an interaction term can be included in the model only if the corresponding main terms are also included in the model. Furthermore, we extend our method to generalized linear models, and show that it performs as well as if the true model were given in advance, that is, it has the oracle property as in Fan and Li (2001) and Fand and Peng (2004). Numerical results on both simulation data and real data indicate that our method tends to remove irrelevant variables more effectively and provide better prediction performance than the classical LASSO method and previous works on this problem.

$\ell_1-\ell_q$ Regularization, Sparse Additive Models, and Simultaneous LASSO

✱Han Liu, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, *hanliu@cs.cmu.edu*; Jian Zhang, Purdue University

Key Words: high-dimensional grouped variable selection, estimation and variable selection consistency, risk consistency, grouped Lasso, Simultaneous Lasso, Sparse Additive Models

We derive the variable section, estimation, and risk consistency results for the high-dimensional least squares regression using $\ell_1-\ell_q$ regularization ($1\leq q\leq \infty$) when the covariates have some predefined group structure. These results provide a unified treatment for the whole family of estimators ranging from Lasso ($q=1$) to iCAP ($q=\infty$), with grouped Lasso ($q=2$) as a special case. Using these results, we also provide theoretical analysis for the sparse nonparametric additive models and the simultaneous Lasso for joint feature selection in the multitask learning setup. We are especially interested in characterizing the conditions under which the $\ell_1-\ell_q$ regularization can provide a better solution than the Lasso estimation. Both simulations and empirical results are provided to illustrate our results.

LARS/LASSO for Feature Selection and Kernel Basis Selection in Machine Learning

✱Sathiya K. Selvaraj, Yahoo! Research, 2821 Mission College Blvd., Santa Clara, CA 95054, *selvarak@yahoo-inc.com*

Key Words: LARS, LASSO, machine learning, feature selection, kernel method

This talk will focus on two applications of LARS/LASSO methods in machine learning: feature selection, and kernel basis selection. In large data mining problems such as text classification that involve a very large number of features, these methods are particularly useful when the number of essential features is less than a few thousand. They offer a much more effective solution compared to filter methods. In the non-linear kernel setting the methods offer a nice way of selecting a small set of basis functions. In this setting these methods will be compared against matching pursuit methods. Both these applications are analyzed in a large scale setting involving a large number of training examples.

Some Extensions of the LASSO: Applications and Algorithms

✱Suhrid Balakrishnan, AT&T Labs - Research, 180 Park Avenue, Flroham Park, NJ 07932, *suhrid@research.att.com*

Key Words: lasso, fused LASSO, online learning

The LASSO (Tibshirani, 1996) has had great success as a practical and theoretical tool for regression and classification problems. We present a few extensions of the Lasso to problems involving: 1. ordered attributes (for example, where input patterns are a set of samples from time series variables). In such cases, identifying and building models involving predictive runs of the attributes leads to highly interpretable models (a variant of the Group LASSO, Yuan and Lin, 2006.) 2. the data arriving in an online fashion (and so cannot be stored in memory). These problems are motivated from applications in telecommunications and epidemiology.

436 Teaching Ethics in Statistics Class

Section on Statistical Education
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Ethics and the Introductory Statistics Course

✱George P. McCabe, Purdue University, 150 N University Street, West Lafayette, IN 47907-2067, *mccabe@stat.purdue.edu*

Key Words: ethics, introductory course

Recent trends in education have recognized the need for inclusion of material on ethics in the undergraduate curriculum. The introductory statistics course provides an ideal opportunity for fulfilling this need. Although other courses are more suitable for providing the broad conceptual framework for the study of ethics, the statistical setting provides a unique opportunity to motivate the study of ethics as a set of principles for guiding decisions. Writing and communication skills can also be developed through assignments and discussion of ethical issues. Some examples of material that has been used in an undergraduate course will be discussed.

Ethics: It's for Everyone

✱Patricia B. Humphrey, Georgia Southern University, Department of Mathematical Sciences, PO Box 8093, Statesboro, GA 30460-8093, *phumphre@georgiasouthern.edu*

Key Words: Ethics, Experiments, Education

Data ethics (or lack of them) hit the headlines periodically: a Korean researcher faked data on cloning; students in California made up answers to a survey about moving a controversial trial; pharmaceutical companies have been accused of withholding information on dangerous side effects. Students "get" that certain experiments would be unethical, but what about other considerations? What should be the role of ethical considerations in a basic statistics course? Texts are starting to incorporate explicit segments on the subject of ethics; discussions on ethics hopefully, are becoming more common. I'll discuss my views on the subject, as well as give some examples of ethical lapses I have encountered, and how I dealt with them.

Truth, Damn Truth, and Statistics

✱Paul F. Velleman, Cornell University, Ives Hall, Ithaca, NY 14850, *pfv2@cornell.edu*

Key Words: lies, damn lies, Twain

Statisticians and Statistics teachers often have to push back against the popular impression that Statistics teaches how to lie with data. Those who believe incorrectly that Statistics is a branch of Mathematics, often see the use of judgment in Statistics as evidence that we manipulate our results. In the push to teach formulas and definitions, we may fail to emphasize the role of judgment. We should teach our students that they are personally responsible for the judgments they make. But we must also offer guidance for their statistical judgments. The ideal guiding these judgments should be the honest search for truth about the world, and the goal of seeking such truth should have a central place in Statistics courses.

437 Mode Effects ●

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Mode Effects on In-Person and Internet Surveys: A Comparison of the General Social Survey and Knowledge Networks Surveys

✱Tom W. Smith, National Opinion Research Center, 1155 East 60th St., Chicago, IL 60637, *smitht@norc.uchicago.edu*; J. Michael Dennis, Knowledge Networks

Key Words: Mode, survey research

A series of experiments have been carried comparing the 2002–2006 General Social Survey with data collected from the Web-enabled Knowledge Networks Panel surveys. First, the results indicate that the level of don't knows are highly contingent on format and layout. It is possible however to design in-person and web surveys to produce similar and comparable levels of item on-response. Second, the substantive distributions are not statistically different across modes for the majority of items. Third, statistically significant and substantively large mode effects do appear for an appreciable minority of items. These differences probably relate to the different demand characteristics

Comparing Estimates and Data Quality from the School Health Policies and Programs Study (SHPPS) 2006 for Mail and Telephone Data Collection: District-Level Questionnaires

✱Maxine Denniston, Centers for Disease Control and Prevention, 2900 Woodcock Blvd, Mail Stop K-33, Atlanta, GA 30341, *mmd1@cdc.gov*; Nancy Brener, Centers for Disease Control and Prevention

Key Words: surveys, data quality, mode effect

SHPPS 2006 collected data on school health policies and programs from nationally representative samples of states, districts, schools, and classrooms. District-level questionnaires were designed for telephone administration, but about half of respondents completed and returned paper questionnaires. Chi-square tests were used to test for differences in population estimates obtained from mail vs. telephone data collection and in the proportions of missing and don't know responses. No strong evidence that population estimates were affected by mode of response was found for any of the 7 questionnaires. As expected, the percent of missing data was higher for mail and the use of don't know was higher for telephone. For example, for the Health Education questionnaire, 3.3% (mail) and 0.0% (telephone) of responses were missing whereas the percentage of don't know responses were 0.1% and 2.4%.

Assessing Mode Bias in a Mixed Mode Reproductive Health Survey

✱Holly B. Shulman, Centers for Disease Control and Prevention, 205 College Ave, Swarthmore, PA 19081, *hbs1@cdc.gov*; Nedra Whitehead, Research Triangle Institute

Key Words: mode effects, mode bias, surveillance, mixed mode

PRAMS is a mixed mode (mail/telephone), state population-based surveillance system of women who recently delivered a live-born infant. Because of the sensitive nature of some of the topics addressed, we examined the extent to which the mode of completion biased survey estimates. We selected a nested matched case-control sample (n=8617) from all respondents, with telephone respondents treated as cases and mail respondents as controls. Matching variables were related to the likelihood of response and to mode of response. We examined 20 sensitive questions. Several socially undesirable events were reported more frequently by mail respondents. The magnitude of the bias ranged from 0.03 to 1.04 percentage points for the 6 events with the strongest mode effects. Our results indicate that mode effects do produce bias in some prevalence estimates from PRAMS, but the magnitude is small.

2006 Canadian Census Internet Mode Effect Study

✱Chantal Grondin, Statistics Canada, 100, Tunney's Pasture Driveway, 15th floor, section N, Ottawa, ON K1A 0T6 Canada, *Chantal.Grondin@statcan.ca*; Limei Sun, Statistics Canada

Key Words: Mode effect, Propensity score, Internet collection, Census

For the first time in 2006, Canadian households had the option of responding to the Census via the Internet. Almost one in five households chose to report their data with this new collection mode. A study was undertaken to check for the presence of an Internet mode effect in the data. Different means were used to assess the presence of a mode effect: non-response rates to each question by response mode were looked at, as well as answers distributions. Among other means, the propensity score method was used to compare answer distributions. This method allowed us to take into account some of the differences in characteristics of Internet reporters and Paper reporters and hence make them more comparable. The findings of the 2006 Census Internet Mode Effect Study are given in this paper.

Evaluating Frames and Modes of Contact in a Study of Individuals with Disabilities

✱Virginia M. Lesser, Oregon State University, Department of Statistics, 44 Kidder Hall, Corvallis, OR 97330, *lesser@science.oregonstate.edu*; Lydia Newton, Survey Research Center; Danny Yang, Survey Research Center

Key Words: survey mode, disabilities, frame

In sampling rare populations, it is expensive to obtain individuals with the characteristic of interest from a random digit dialing survey. In a study of individuals with a disability, we obtained opinions about their use of air travel. Two approaches were used to select a sample. One approach used a list of individuals that are members of a disability association. A random sample of these members was mailed a questionnaire and opinions about air travel was collected. A second but more expensive approach used random digit dialing methods to obtain a sample of individuals with a disability. In both approaches, once an individual with a disability who traveled by air was recruited for the study, they were randomized to complete a questionnaire using telephone, mail or web. The results obtained from the two frames and for the different modes will be presented.

438 Modeling of Migration and Social Networks ●▲

Social Statistics Section, Section on Government Statistics, Section on Statisticians in Defense and National Security

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Estimation of International Migration Flow Tables

✱Guy J. Abel, University of Southampton, School of Social Sciences, Southampton, International SO17 1BJ United Kingdom, *g.j.abel@soton.ac.uk*

Key Words: EM Algorithm, International Migration, Missing Data

Problems of inconsistencies and missingness in international migration flow data are widespread. Such problems restrict the construction of reliable migration flow tables desired by regional policy makers and social scientist alike. Previous attempts to harmonize flow data have concentrated on ad-hoc scaling methods to enable the construction of complete flow tables. In this presentation, for reported values of movements between Northern European nations, a model-based approach is taken to examine known reliable data. Firstly, good quality data are adjusted to a harmonized definition. The Expectation-Maximization (EM) algorithm is then employed to impute missing flow values. This enables the construction of a complete table of harmonized flows. Finally, measures of variability on the new estimates are then derived using the Supplemented EM algorithm.

Statistical Modeling of Migration Flows

✱Peter W.F. Smith, University of Southampton, S3RI, School of Social Sciences, University of Southampton, Southampton, International SO17 1BJ United Kingdom, *P.W.Smith@soton.ac.uk*; James Raymer, University of Southampton; Corrado Giulietti, University of Southampton

Key Words: internal migration, combining data, England and Wales, log-linear modelling

In England and Wales, there are two main sources of internal migration data: censuses and health service registers. The censuses provide detailed data but only once every ten years. The health service registers provide annual data but with less detail. We develop a log-linear model for combining the two data sources to predict detailed migration flows over time from 2001 to 2007. In our illustration, the detailed flows are the migration counts cross-classified by origin, destination, age and sex, and ethnicity, which is only available from the decennial censuses. The result is a synthetic and consistent data base that can be used, say, by local governments to improve their planning policies directed at supplying particular social services or at influencing levels of migration.

Separable Models for Dynamic Social Networks

✱Pavel N. Krivitsky, University of Washington, Department of Statistics, Box 354322, Seattle, WA 98195-4322, *pavel@stat.washington.edu*; Mark S. Handcock, University of Washington

Key Words: dynamic networks, exponential random graph models, method of moments

Models of dynamic networks—social networks that evolve over time—have manifold applications in social sciences, epidemiology, and other situations where evolving relational data are observed. We develop and apply a discrete-time generative model for social network evolution as a series of formation and dissolution steps, generated from conditional exponential-family random graph models (ERGMs). This model inherits the richness and flexibility of ERGMs and facilitates modeling of tie duration distributions, while avoiding pitfalls of some of the past efforts in this area. Method of Moments is used to fit this model to observed cross-sectional and egocentric networks and incorporate available tie duration data.

Using Social Space Models for Inference on Missing Nodes

✱David Marchette, Naval Surface Warfare Center, 18444 Frontage Rd. Suite 327, Code Q21, Dahlgren, VA 22448, *dmarchette@gmail.com*

Key Words: random dot product graph, social network analysis, latent variables

Social network analysis generally assumes a model for the network giving the probabilistic edge structure of the graph. One class of models posits a "social space," defined by latent variables, where the position of the vertices in the space determines the edge probabilities. This has been used for various inferences including predicting missing edges and assigning labels to unlabeled vertices. Often one assumes that all the vertices have been observed, or that the only vertices about which information is sought are those that have been observed. It may seem that the very nature of these models make it impossible to make inferences about missing vertices. We discuss a method for making these inferences, under certain assumptions about the social space, and give some examples of this inference using a simple model which can be fit using a least squares algorithm.

Estimating Enforced Disappearances in Northern India: Comparative Notes on Multiple Systems Estimation and Respondent-Driven Sampling

✱Romesh Silva, University of California, Berkeley, Demography Department, 2232 Piedmont Ave, Berkeley, CA 94720, *romeshs@demog.berkeley.edu*

Key Words: Human Rights, Multiple Systems Estimation, Respondent Driven Sampling, Hidden Populations, Statistical Learning, Record Linkage

Existing claims about the magnitude of enforced disappearances in Punjab (Northern India) between 1984 and 1996 have not been based on defensible statistical estimation processes. This paper reviews and compares the applicability of Multiple Systems Estimation (MSE) and Respondent Driven Sampling (RDS) to the measurement of hidden populations, such as victims of enforced disappearances. By drawing on six independently collected data sources, we derive MSE estimates of the magnitude and pattern of total enforced disappearances in Punjab. These estimates are derived by using record-linkage methods which draw heavily from recent advances in statistical learning. We also present insights into RDS gained from a series of small pilot tests in three districts of Punjab, Northern India. We conclude by comparing the applicability of MSE and RDS in answering policy-relevant human rights questions.

Your House Is Not Just a Roof Over Your Head: Assessing the Relationship Between Neighborhood Stability and Home-Ownership

✱Germana Corrado, University of Rome Tor Vergata, via Columbia 2, Rome, 00133 Italy, *corrado@uniroma2.it*; Luisa Corrado, University of Cambridge

Key Words: Spatial Effects, Hierarchical Modelling, Housing Economics

The relationship between homeownership and neighborhood stability is taken for granted by housing practitioners since homeowners are believed to be more involved in civic affairs and to maintain houses at higher standards. This, in turn, leads to more stable neighborhood and higher property values. Using the British Household Panel Survey we assess whether homeownership and other intermediate home-buying schemes such as shared-ownership can really produce such virtuous effects. The principal issue is that neighborhood stability is determined by a combination of individual characteristics of the homeowner, together with the influence of the household structure, residence, peer groups and other groupings effects. As we have data organized by area of residence we use a hierarchical modeling approach.

439 Adjusting for and Minimizing Unit Nonresponse ●

Section on Survey Research Methods, Section on Government Statistics

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Update on Use of Administrative Data To Explore Effect of Establishment Nonresponse Adjustment on the National Compensation Survey Estimates

Chester H. Ponikowski, Bureau of Labor Statistics; Erin McNulty, Bureau of Labor Statistics; ✱Jackson Crockett, Bureau of Labor Statistics, 2 Massachusetts Ave., NE, Room 3160, Washington, DC 20212, *crockett.jackson@bls.gov*

Key Words: non-response, bias, weighting cells

A common technique used to reduce nonresponse bias is to adjust the sampling weights of responding units to account for nonresponding units within a set of weighting cells. In the National Compensation Survey (NCS), which is an establishment survey, the weighting cells are formed using available auxiliary information: ownership, industry, and establishment size. At JSM 2006, we presented a paper in which we explored the effectiveness of the formed cells in reducing potential bias in the NCS estimates and presented results for one NCS area. Since 2006, NCS has expanded this study to several additional survey areas and time periods. In this paper, we present results from this additional research. We include areas of different size and with different levels of nonresponse. Also we compare the direction and magnitude of bias across time and across areas.

Empirical Evaluation of Raking Ratio Adjustments for Nonresponse

✱Ismael Flores Cervantes, Westat, 1650 Research Blvd., Rockville, MD 20850, *ismaelflores-cervantes@westat.com*; J. Michael Brick, Westat

Raking ratio adjustments are generally used to benchmark sampling weights to known control totals as a form for face validity. Raking is also used to reduce sampling error thorough the use of auxiliary variables correlated to survey response (Deville and Särndal, 1992). Lundström and Särndal (1999) shows that raking can be used also as a method for treatment of nonresponse reducing both sampling error and sampling error and non-response bias. The use of raking can expedite the creation of analysis weights by simplifying the number of weighting adjustments. In this paper, we empirically evaluate analysis weights created by raking sampling weights not adjusted for nonresponse. The raked based weights are compared with analysis weights created by applying sequential weighting adjustments at each stage of nonresponse in a survey.

The Bitter End? The Close of the 2007 SCF Field Period

✱Arthur B. Kennickell, Federal Reserve Board of Governors, Mail Stop 153, Washington, DC 20551, *Arthur.Kennickell@frb.gov*

Key Words: Unit nonresponse, Contact strategy, Survey of Consumer Finances

Typically, the end of the field period is a point when a survey struggles to achieve what is seen as an acceptable response rate. Response is shaped by the decisions of field staff to continue applying effort and respondents to be open to persuasion. The situation is an uncomfortable one in that it is quite difficult to apply measurable standards to the process in a way that can be mapped with any precision into response probabilities. The Survey of Consumer Finances uses a phased contacting strategy to provide a measurable framework for most of field effort, but the terminal phase is open ended. This paper attempts to use information from the progress of cases through the first contact phases to clarify the nature of the residual set of completed interview. It also draws on call records to classify the cases in the terminal phase by the level of resistance.

Identifying, Collecting, and Using Auxiliary Variables To Adjust for Nonresponse Bias in Organizational Surveys

✱Ashley Bowers, The University of Michigan, *afbowers@umich.edu*

Key Words: nonresponse bias, organizational surveys, weighting adjustment

With low response rates of 20-30% and evidence of differences between respondents and nonrespondents, it is clear that nonresponse bias is a problem that plagues many organizational surveys, yet the problem remains largely unaddressed (Assael & Keon, 1982; Cycyota & Harrison, 2006; Paxson, 2002; Smith, 1997). Nonresponse bias is a function of (a) the likelihood of responding, and (b) the correlation between the variable of interest and the likelihood of responding. One strategy to reduce nonresponse bias is to collect auxiliary variables (Zs) that are associated with the variable of interest (Y) and the propensity to respond (P) and to use them in a weighting adjustment (Groves, 2006). In this paper, we demonstrate the use of auxiliary variables in several post-survey adjustments for the Community Health Measures Survey and propose future work on the use of Zs in organizational surveys.

Incorporating Multiple Reasons for Attrition into Analysis of Longitudinal Data

✱Irina Bondarenko, The University of Michigan, *ibond@umich.edu*; Trivellore Raghunathan, The University of Michigan

Key Words: Missing data, nonignorable mechanism, attrition, longitudinal, multiple imputation

In many longitudinal studies, subjects may drop out due to a variety of reasons. For example, in a study monitoring the post-surgery progress of infants suffering from Biliary atresia, a serious disease in which the ducts that carry bile from the liver to the gallbladder are injured or underdeveloped, subject may drop out due to death, liver transplant or loss-to-follow up. The distribution of responses across the drop out groups could vary systematically and the missing data mechanism may be nonignorable for some of them. We develop a multiple imputation method where the completers are matched to subjects with a particular reason for drop out based on covariates and the observed outcome and then constrain the imputation of the missing values to matched subjects. We illustrate this methodology using the data from Biliary Atresia Research Consortium Network. Simulation study is presented.

Subjective and Objective Numeracy and Disclosure Risk in Surveys

✱Mick P. Couper, The University of Michigan, *mcouper@umich.edu*; Eleanor Singer, The University of Michigan

Key Words: risk disclosure, informed consent, numeracy, survey participation

A web-based study (n=6400) with hypothetical vignettes was used to explore willingness to participate (WTP) in surveys, varying disclosure risk, topic sensitivity, & mode. We also asked respondents for their preferences for risk communicated in words or in numbers, & measured objective and subjective numeracy. We present 3 sets of analyses: 1) relationship between objective & subjective numeracy, & correlates of numeracy, 2) how numeracy relates to preference for words or numbers in describing risk, & 3) effect of numeracy on stated WTP given the risk manipulation. Numeracy relates to preferences for risk description, & interacts with the risk manipulation in predicting WTP: the effect of the risk manipulation is larger for those with higher levels of numeracy & for those who prefer numbers over words. The findings have implications for disclosure risk communication in surveys.

440 Improved Sample Designs ●

Section on Survey Research Methods, Section on Government Statistics

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Overview of the Survey of Occupational Injuries and Illnesses Sample Design and Estimation Methodology

✱Philip N. Selby, Bureau of Labor Statistics, 2 Massachusetts Ave NE, Room 3160, Washington, DC 20212, *Selby.Philip@bls.gov*; Terry M. Burdette, Bureau of Labor Statistics; Erin Huband, Bureau of Labor Statistics

Key Words: stratified sample, target estimation industry

The Survey of Occupational Injuries and Illnesses (SOII) uses a stratified sample to produce State and National estimates for nonfatal workplace injuries and illnesses. Private industry estimates are produced separately for 43 States, the District of Columbia, and three U.S. territories (Guam, Puerto Rico and the Virgin Islands). The level of industry detail for which State estimates are available varies widely and is based on the needs determined by each State. Additionally, estimates for injuries and illnesses for State and Local Government workers are available for 26 of these States. This paper will describe the frame development, target estimation industry identification, sample allocation, sample selection, and estimation methodologies used to produce the number and frequency (incidence rates) of nonfatal workplace injuries and illnesses.

Evaluating the Within-Household Selection Procedures for in-Person U.S. Adult Literacy Surveys

✱Tom Krenzke, Westat, 1650 Research Blvd, Rockville, MD 20850, *TomKrenzke@Westat.com*; Leyla Mohadjer, Westat; Lin Li, Westat

The 2003 National Assessment of Adult Literacy (NAAL) and the international Adult Literacy and Lifeskills (ALL) surveys each involved stratified multistage area sample designs. During the last stage, a household roster was constructed, the eligibility status of each individual was determined, and the selection procedure was invoked to randomly select one or two eligible persons within the household. The objective of this paper is to evaluate the within-household selection procedure used in the 2003 surveys and update the procedure for future literacy surveys. The analysis will be based on current household composition data, and intracluster correlation coefficients for covariates related to literacy. In our evaluation, several feasible household selection rules are studied, considering effects from clustering, differential sampling rates, cost per interview, and household burden.

Developing Statistical 'Twins' Methodology for Selecting Sites for Qualitative Case Studies

✱Zhiwei Zhang, National Opinion Research Center, 4350 East West Hwy, Suite 800, Bethesda, MD 20814, *zhangz2007@gmail.com*; Fritz Scheuren, The University of Chicago

Key Words: survey methodology, site selection, case study, similarity, dissimilarity, distance matrix

Site selections for qualitative substance abuse evaluation studies are usually constrained due to cost concerns not only by the limited number of sites selected but also by the casual and subjective method for selecting particular sites. A methodology aimed at identifying twin-counties with similar demographics yet different substance use and mental health profiles is developed. Social distance matrices are constructed to list the degrees of similarity among all possible pairs of counties within each state based on predetermined source variables. The distance matrices are obtained separately through socio-demographic characteristics and through the substance abuse, mental health and service coverage measures. Two ranking indexes are created and composite scales for ranking are established. The methodology is further assessed with qualitative case study result from multipaired twin sites.

Sample Design and Sample Selection for Adult Multivitamin Mineral Study

✱Charles Perry, USDA ARS NDL, 12402 Washington Brice Rd, Fairfax, WA 22033, *cperry7@cox.net*; Joanne Holden, USDA ARS NDL; Janet Roseland , USDA ARS NDL; Larry Douglas,

Key Words: multivitamin minerals, average daily dose, nutrient content, representative sample

The goal of the study was to generate national level estimates of the nutrient content of average daily dose of adult multivitamin minerals (AMVM) sold in the United States from a representative sample of the products from each of the groups: 1. For the top 16 brands, the objective was to obtain estimates of the nutrient content for 35 of the most important products. This group accounts for approximately 86% of the total market-share of all AMVM. 2. For the next 12 brands, the objective was to obtain estimates of the nutrient content for the most representative product. This group accounts for approximately 7% of total market-share. 3. For all other brands, the objective was to obtain estimates of the nutrient content for a sample of 15 statistically representative products/brands. The total marketing share for the 77 brands in this group was approximately 7%.

Optimized Whole-Sample Procedures vs. Traditional Draw-by-Draw Procedures

✱Sun-Woong Kim, Dongguk University, Department of Statistics, 26, 3 Ga Pil-Dong, Chung-Gu, Seoul, International 100-715 South Korea, *sunwk@dongguk.edu*; Steven G. Heeringa, Institute for Social Research; Sung-Joon Hong, Dongguk University; Peter S. Solenberger, Institute for Social Research

Key Words: Superpopulation model, Expected variance, Joint probabilities

Kim, Heeringa, and Solenberger (2006) suggested model-based sampling methods for reducing the variance of the Horvitz-Thompson (1952)'s estimator. Their methods are whole sample procedures based on optimization theory under a superpopulation model. With respect to the sample selection probabilities, we theoretically present the differences between those procedures and popular draw-by-draw procedures such as sampling methods of Mizuno (1952), Brewer (1963), and Murthy (1957), when the sample size is two, which is a common situation in nationwide samples with many strata. We also compare the efficiencies between them for natural populations in the published literature. In addition, we empirically compare the properties of the sample selection probabilities obtained from the procedures through a graphical display.

Intraclass Correlation Patterns of Cognitive and Behavioral Measures of Illicit Drug Use Within Six Major Metropolitan Areas

Zhiwei Zhang, National Opinion Research Center; ✱Douglas Wright, National Opinion Research Center, 4350 East West Hwy, Suite 800, Bethesda, MD 20814, *douglaswright131@gmail.com*

Key Words: Intraclass correlation, drug use, NHSDA, NSDUH, prevention, intervention

Through the National Household Survey on Drug Abuse (the predecessor of the current National Survey on Drug Use and Health), we calculated intraclass correlations (IC) for both cognitive and behavioral measures on drug use at the census tract and census block group levels within six major metropolitan areas respectively. We demonstrate that IC should not be overlooked in the substance abuse fields and discuss further the utility of empirical knowledge of intraclass correlations of pertinent measures in future sample design, estimations, and for policy oriented preventions and interventions.

Drawing a Sample from a Given Distribution

✱Andrew Vogt, Georgetown University, Department of Mathematics, Georgetown University, Washington, D. C., DC 20057-1233, *vogt@math.georgetown.edu*; Dhirendra N. Ghosh, Synectics for Management Decisions Inc.

Key Words: Rejective sampling, Inversive sampling, Lahiri's method, CDF, Simulation

Four methods of sampling from any given distribution are considered: natural, inversive, rejective (especially, Lahiri's method), and symmetry-based. In natural sampling equally-likely sampling is done on a finite population that obeys the distribution approximately. Inversive sampling refers to choosing a random sample from a uniform distribution and applying the inverse of the cumulative distribution function (cdf), and to higher-dimensional generalizations. Rejective sampling refers to choosing a random sample from a product of independent uniform distributions and rejecting units that violate preimposed constraints. Lahiri's method, in particular, avoids computation of integrals and inverses. Symmetry-based sampling exploits symmetry to transform the sampling into rejective sampling and/or uniform sampling. One application is equal area/volume sampling from a geometric surface/body.

441 Measurement Error and Control of Bias ●

Section on Statistics in Epidemiology, Section on Survey Research Methods, Social Statistics Section, Biometrics Section

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Monitoring Quality Control: Can We Get Better Data?

✱Ofer Harel, University of Connecticut, 215 Glenbrook Road U-4120, Storrs, CT 06269-4120, *oharel@stat.uconn.edu*; Enrique F. Schisterman, National Institute of Child Health and Human Development; Albert Vexler, National Institute of Child Health and Human Development; Marcus Ruopp, National Institute of Child Health and Human Development

Key Words: quality control, epidemiological studies, data collection

Quality control is important in many fields, especially in industry. We adapt and develop methodology in quality control to monitor data collection in epidemiological studies. To date, we do not know of any procedures to evaluate quality control during the actual process of data collection but only after the data has been collected. In this regard, we focus on two important processes during data collection: instrument calibration and population sampling. To evaluate instrument calibration, we present methods utilizing Shewhart control charts and Westgard stopping rules. To evaluate population sampling, we present methods utilizing regression analysis. The proposed methodology is beneficial to investigators to help assess the quality of data they are collecting and to allow them to adjust for data collected that is of low-quality during real-time data collection

Measurement Error Adjusted Effects of A-Bomb Radiation Dose on Cancer Risk

✱Carmen D. Tekwe, University at Buffalo, 249 Farber Hall, 3435 Main Street, Bldg 26, Buffalo, NY 14214-3000, *cdtekwe@buffalo.edu*; Randolph L. Carter, University at Buffalo; Austin Miller, University at Buffalo; Harry Cullings, Radiation Effects Research Foundation; John Cologne, Radiation Effects Research Foundation; Yoichiro Kusunoki, Radiation Effects Research Foundation; Kazuo Neriishi, Radiation Effects Research Foundation; Sachiyo Funamoto, Radiation Effects Research Foundation; Norman P. Ross, Radiation Effects Research Foundation

Key Words: measurement error, non-linear MIMIC, structural equation models, A-bomb survivors, radiation effects, cancer mortality rate

Multiple indicators, multiple causes (MIMIC) models are useful for studying the effects of a latent variable on several outcomes, when causes of the latent variable are observed. The error in the causal equation of a MIMIC model is a Berkson error and the classical Berkson error model is a MIMIC model. Previous work has focused on linear MIMIC models, where the causes of the latent variable are observed without error. We generalize the MIMIC model to allow non-linear relationships and to allow classical measurement error in the explanatory variables of the causal equation. We propose estimation procedures to estimate the resulting G-MIMIC ME model parameters based on the Monte Carlo EM algorithm. We apply our methods to data collected on A-bomb survivors to estimate the effect of radiation dose on cancer mortality, adjusting for both classical and Berkson measurement errors.

Assessing the Impact of Measurement Error in Modeling Change: A Sensitivity Analysis Approach

✱David Yanez, University of Washington, Dept. of Biostatistics, Box 357232, Seattle, WA 98195, *yanez@u.washington.edu*; Thomas Lumley, University of Washington; Chengcheng Hu, Harvard University; Mose Andre, University of Washington

Key Words: method of moments, sensitivity analysis, reliability, errors in variables

This paper investigates the problem of modeling change in some imprecisely measured outcome and precisely measured predictors in the absence of auxiliary data (e.g., replicates, validation data). For models where adjusting for the baseline outcome as a covariate is appropriate, a measurement error induced bias will likely result. Ordinary regression analysis cannot be used to account for measurement error biases; standard measurement error modeling techniques cannot be implemented without auxiliary data. We present a method to investigate associations between change and predictors by parameterizing large sample estimators as functions of the reliability or the measurement error variance and employing sensitivity analysis. This approach can be implemented provided acceptable bounds for the reliability or measurement error variance are specified. An illustration is provided.

Modeling Biases in Observational Data Using Bayesian Graphical Models To Combine Multiple Data Sources: Application to Low Birth-Weight and Water Disinfection Byproducts

✱NuooTing Molitor, Imperial College, London, Division of Epidemiology, Public Health and Primary Care, Research Associate, London, W2 1PG United Kingdom, *jassy.molitor@imperial.ac.uk*; Chris Jackson, MRC Biostatistics Unit; Nicky Best, Imperial College, London; Sylvia Richardson, Imperial College, London

Key Words: Bayesian statistics, multiple sources of data, low birth weight, observational study, MCMC, graphical modeling

Data in the social, behavioral and health sciences frequently come from observational studies instead of controlled experiments. Observational data are filled with different sources of uncertainty such as missing values, unmeasured confounders, and selection biases. As such, a single data set may not provide all the necessary information, so multiple data sources are often required to identify biases and inform about different aspects of the research question. In this manuscript, we present a unified Bayesian modeling framework that will account for multiple biases simultaneously and give more accurate parameter estimates than standard approaches. We illustrate our approach by analyzing data from a study of water disinfection byproducts and adverse birth outcomes in the U.K.

Bounds on ACE and Unmeasured Confounding Bias in Observational Studies

✱Tao Liu, Brown University, 46 Roger Williams Green, Providence, RI 02904, *tliu@stat.brown.edu*; Joseph Hogan, Brown University; Allison DeLong, Brown University

Key Words: Confounding, Observational study, Instrumental variable, Causal effect, Bounds

We consider two problems in observational studies when unmeasured confounding exists: The estimation of average causal effect (ACE) of an exposure and the quantification of unmeasured confounding. We use an instrumental variables (IV) method, combined with contextually appropriate assumptions, to place bounds on ACE. We account for measured confounding using inverse probability weighting (IPW) and incorporate a sensitivity parameter into the IPW estimating equations to encode the unmeasured confounding. The bound on ACE from IV allows us to infer the plausible range of the sensitivity parameter, and hence the magnitude of unmeasured confounding. We apply our methods to the HIV Epidemiology Research Study, where one primary interest is the initial stage ACE of highly active antiretroviral therapy on CD4 count among HIV+ women, and the existence of unmeasured confounding is of concern.

Misclassification Adjustment in Threshold Models for the Effects of Subject-Specific Exposure Means and Variances

✱Chengxing Lu, Emory University, 3537 Chattahoochee Summit Ln., Biostatistics Dept., Atlanta, GA 30339, *clu@sph.emory.edu*; Robert H. Lyles, Emory University

Key Words: Threshold Model, Misclassification, Repeated Measurements, Random Effects

In environmental epidemiology, researchers sometimes assume the existence of exposure thresholds above which the risk of adverse effects begins to increment. In this work, the research question of interest is to identify the relationship between a health-related outcome and whether or not the subject-specific mean and/or variability of the exposure exceed known thresholds. As a subject's true exposure mean and variability cannot be observed directly, misclassifications usually arise. Building off of random effects models for repeated exposure measurements and assuming balanced data, methods based on regression calibration and matrix methods are demonstrated for mean exposure only. For unbalanced data, and to incorporate categorizations based on both exposure mean and variance, a maximum likelihood approach is introduced. Simulation results and a real study example are also presented.

Statistical Methods for Biodosimetry in the Presence of Both Classical and Berkson Measurement Error

✱Austin Miller, University at Buffalo, 249 Farber Hall, Department of Biostatistics, 3435 Main St. Bldg. 26, Buffalo, NY 14214-3000, *am65@buffalo.edu*; Randolph L. Carter, University at Buffalo; Harry Cullings, Radiation Effects Research Foundation; John Cologne, Radiation Effects Research Foundation; Norman P. Ross, Radiation Effects Research Foundation; Kazuo Neriishi, Radiation Effects Research Foundation; Carmen D. Tekwe, University at Buffalo

Key Words: Biodosimetry, Measurement Error, Berkson Error

In radiation epidemiology cases, the true dose received by those exposed cannot be assessed directly. Physical dosimetry uses a deterministic function of the source term, distance and shielding to estimate dose. Biological dosimetry inverts a well-defined dose-response curve at a biological indicator of dose. For the Atomic Bomb survivors, the physical dosimeter system is well established, and biological indicators have recently come available. We propose a likelihood-based dose-estimation method that incorporates both the physical and biological indicators. The classical measurement error plaguing the location and shielding data supporting physical dosimetry system is well known. More recently, researchers have realized the added complexity of Berkson error in the dose estimates. Our proposed method accounts for both error types.

442 Pollution and Health

Section on Statistics and the Environment, Section on Statistics in Epidemiology, Section on Health Policy Statistics

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

The Effect of PM Exposure on Birth Weight

✱Simone Gray, Duke University, Nicholas School of the Environment / Earth Science, Box 90328 - Duke University, Durham, NC 27708-0328, *simone@stat.duke.edu*; Alan Gelfand, Duke University; Marie Lynn Miranda, Duke University; Geeta Swamy, Duke University; Sharon Edwards, Duke University

Key Words: spatial epidemiology, air pollution, birthweight

Studies suggest air pollution exposure may be associated with adverse birth outcomes. It is challenging to assess pollution exposure during pregnancy. How do we introduce exposure into an explanatory model for birth weight? Should it be cumulative, episodic, extremes, exceedances, overall or by trimester? In this talk we discuss a variety of modeling strategies used to better understand these issues. Moreover, if we consider a large collection of birth weights with geocoded maternal residences, we would anticipate spatial dependence in exposure. How can we introduce spatial structure in a suitable fashion for this situation? How do we accommodate necessary spatial modeling with more than 100,000 births? We utilize the North Carolina Detailed Birth Records database, linked with the PM10 and PM2.5 monitoring data from the USEPA to provide preliminary answers to some of these questions.

Informative Priors for Computing the UCL of a Set of Asbestos Measurements with Low Total Counts

✱Bradley C. Venner, USEPA, PO Box 25227, Denver, CO 80232, *venner.brad@epa.gov*; Bill Brattin, Syracuse Research Corporation; Terry Schulz, Independent Contractor; Timothy Barry, USEPA

Key Words: asbestos, Poisson-lognormal, exposure point concentration, Superfund

USEPA guidance on risk assessment at Superfund sites recommends calculation of an exposure point concentration as an upper confidence limit on the mean of the samples. This approach is difficult when measuring asbestos. These measurements often occur with few if any observed fibers due to the cost of manually counting asbestos structures under high magnification. Another complication is that there are two major sources of variability—in the observed number of structures due to counting errors and the air concentration of interest. Although a compound Poisson-lognormal model may be used to describe both sources of variability, this model suffers from a lack of identifiability at low number of counts. We propose the use of informative priors for the log-normal parameters, present the results of simulation studies on the sensitivity of this prior, and discuss experimental design.

Imputation of the Physical Activity Component of Gerontologic Frailty for Use in Modeling the Health Effects of Air Pollution on Older Adults in the Cardiovascular Health Study

✱Sandrah P. Eckel, Johns Hopkins Bloomberg School of Public Health, Department of Biostatistics, 615 N Wolfe St, Room E3527, Baltimore, MD 21205, *seckel@jhsph.edu*; Thomas A. Louis, Johns Hopkins University; Karen Bandeen-Roche, Johns Hopkins Bloomberg School of Public Health; Paulo Chaves, Johns Hopkins Bloomberg School of Public Health; Linda P. Fried, Johns Hopkins Medical Institutions; Helene G. Margolis, California Department of Public Health

Key Words: effect modification, air pollution effects, geriatrics, missing data

Our goal is to develop, evaluate and apply statistical methods to assess the hypothesis that the health effects of air pollution are modified by underlying gerontologic frailty status as well as age and comorbidity. We use data from the Cardiovascular Health Study (CHS) and the CHS Environmental Factors Ancillary Study. In gerontology, frail individuals are characterized by slow walking speed, exhaustion, unintended weight loss, weakness and low physical activity. In CHS, physical activity was measured only at a subset of visits. To fill in missing values, we develop a longitudinal frailty status variable using surrogate measures of physical activity to impute low Kcal/week expenditure for the missing values. We report on the imputation method and on initial analyses of the modifying effect of frailty on air pollution and health associations.

Developing New Methodologies To Estimate the Proportion of Homes in a Region with Indoor Radon Levels Exceeding a Reference Threshold

✱Patrick E. Murphy, University College Dublin, School of Mathematical Sciences, UCD, Belfield, Dublin, International 0 Ireland, *patrick.murphy@ucd.ie*; Catherine Organo, Radiological Protection Institute of Ireland

Key Words: Radon, Mapping, Beta, Modelling

The World Health Organization estimates that between 6% and 15% of all lung cancer deaths per year are caused by exposure to indoor radon. The statistical task involved in producing radon maps is to estimate the proportion of homes, P, in a particular region which exceed a reference level of 200

Bqm3 and consequently require remediation. Early work conducted in this field established that radon levels could be modeled using a Log-Normal distribution. This paper will describe the development of alternative modeling approaches using the Log-Normal, Gamma and Beta distributions to produce estimates for P. We will then report on the results of recent work which investigates the effect of outliers on predictions of P. We also provide for the first time accurate determinations of appropriate sample sizes to be used in surveys for the construction of radon maps.

Modeling Particulate Matter Emissions Indices at the Hartsfield-Jackson Atlanta International Airport

✱Lu Gan, Missouri University of Science and Technology, Department of Mathematics & Statistics, 1870 Miner Circle, Rolla, MO 65409, *lgan@mst.edu*; Gary L. Gadbury, Kansas State University; Prem Lobo, Missouri University of Science and Technology ; Philip D. Whitefield, Missouri University of Science and Technology ; Donald E. Hagen, Missouri University of Science and Technology

Key Words: Particulate Matter, Emissions Indices, Linear Model, Bootstrap, Confidence Intervals, Aircrafts

Particulate Matter (PM) emissions from aircrafts have been identified as possible contributors to global climate and local air quality changes. An objective of this study was to develop a predictive model for emissions indices (EI) of PM emitted from aircraft during normal operation given engine type and ambient conditions. A model that is scientifically “meaningful” was desirable as it could help guide follow-on activities in air quality monitoring at airports. Diagnostics, transformations, and manipulation of some variables suggested a linear model would be appropriate. Temperature, pressure, and a temperature-pressure interaction were key variables. Bootstrap confidence intervals for mean EI were computed since residuals were slightly heavy tailed. We hope this finding can be applied to EI estimation, plume evolution, model development, and local air quality modeling applications.

Assessing Water Pollution Using Lognormal Distribution in a Generalized Linear Modeling Framework

✱Obisesan O. Kazeem, University of Ibadan, Department of Statistics, Ibadan, International +234 Nigeria, *ko.obisesan@mail.ui.edu.ng*

Key Words: Evaluation, Generalised Linear Model, Deviance, Shapiro-Wilk, Lognormal Model

The environment has been polluted by man through the introduction of substances liable to cause hazards to human. Evaluation of pollutants in Eleyele reservoir in Ibadan, was considered by using the lognormal distribution with the objective of identifying pollutants with the highest pollution risk .The Shapiro-Wilks method was used to check the suitability and the goodness-of-fit of the model while the deviance statistic was used for model comparison. The GLM results confirmed that the DO concentration was insignificant (p-values greater than 0.05) and COD was the most significant (p-value less than 0.05). The low level of DO indicate a threat to live for aquatic organisms. The Shapiro-Wilks Model showed that the distribution of pollutants was not normally distributed; hence the lognormal model is a good fit. The deviance values also support the lognormal model.

443 Complex Data Issues in Marketing Research ●

Section on Statistics and Marketing

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

The Application of Poisson Race Model to Conjoint Choice Study with Multiple Alternatives

✱Shiling Ruan, U.S. Food and Drug Administration, Division of Biostatistics, Center for Devices and Rodiological Health, 1350 Piccard Dr, Rockville, MD 20850, *shiling.ruan@fda.hhs.gov*; Steven N. MacEachern, The Ohio State University; Thomas Otter, Goethe University Frankfurt; Angela M. Dean, The Ohio State University

Key Words: Poisson race model, conjoint choice experiment, multinomial logit model, attribute-based process, stochastic process, dependence structure

Poisson race model with a dependence structure has been shown to model the conjoint choice data with two alternatives. Drawing upon the results of the model with two alternatives, this paper discusses the formulation of Poisson race model with a dependence structure for more than two alternatives. This model takes into account the various dominance relationships among the attributes of alternatives and derives an attribute-based choice process with a stochastic model. The model is then applied to real conjoint choice data with a mixture of two or three alternatives in a choice sets. The model parameters are estimated by hierarchical Bayesian methods. The Poisson race model demonstrates its ability to capture interesting choice behaviors with superior predictive performance as compared to traditional choice models including the multinomial logit model.

Predictive CRM in a B2B Context: Model-Based Cross- and Up-Selling Support for Guiding Recommendations and Improving Sales

✱Lynd Bacon, Loma Buena Associates, Ltd., P.O. Box 62960, Woodside, CA 94062-0960, *lbacon@LBA.com*; Peter Lenk, The Stephen M. Ross Business School, University of Michigan; Michael Troutman, C.M. Troutman Marketing Analytics and Consulting; Ashwin Sridhar, ZLINQ Solutions, Inc.

Key Words: CRM, cross-selling, Hierarchical Bayes, recommendation systems, sales support, decision support systems

Most companies struggle with the issue of what products to offer their existing customers. An important CRM goal is offering customer what they are most likely to buy, not necessarily what they they ask for. The key is to offer customers what they are must likely to buy, given that an offer or recommendation is made. We describe the development of a B2B CRM system that provides tactical sales support through the use of sensor (GIS) and account management data. It deploys hierarchical Bayes models of purchase behavior that condition on sales interventions, and that incorporate customer covariates to account for heterogeneity. We compare our procedure to the company’s current practices, and provide an contextual computing IT systems perspective.

Choice Experiments for Estimating Main Effects and Two-Factor Interactions Inclusive of One Factor

✱Pallavi Chitturi, Temple University, 1810 North 13th Street, 341 Speakman Hall, Philadelphia, PA 19122, *chitturi@temple.edu*; Jing Chen, Temple University

Key Words: factorial experiments, choice experiments, pareto optimal subset

Choice-based conjoint experiments are used when choice alternatives can be described in terms of attributes. The objective is to infer the value that respondents attach to attribute levels. This method involves the design of profiles on the basis of attributes specified at certain levels. Respondents are presented sets of profiles and asked to select the profile they consider best. Choice sets with no dominating or dominated profiles are called pareto optimal. However, if choice sets have too many profiles they may be difficult to implement. We provide strategies for reducing the number of profiles in choice sets. We consider situations where only a subset of two factor interactions is of interest. For factorial experiments, we obtain connected main effects plans with smaller choice sets that are capable of estimating all two factor interactions inclusive of one specific factor.

Determining the Impact of Attitudinal Drivers on Customer Satisfaction in the Presence of Missing Data

Kurt A. Pflughoeft, Market Probe; ✱Jorge A. Alejandro, Market Probe, 2655 N Mayfair Rd, Milwaukee, WI 53226, *j.alejandro@marketprobe.com*

Key Words: Missing Data, Multiple Imputation, Attitudinal Drivers, Relative Importance, Random Forest, MARS

With the advent of the Multi-Attribute Attitudinal Model in the early 1960's, researchers have been concerned not only with measuring attitudes related to customers satisfaction but also with determining the importance of each attribute. Estimating the importance of attitudes on overall satisfaction is a vexing problem and missing values can only further exacerbate it. Multiple imputation is used in conjunction with several importance techniques which measure either predictive accuracy or overall information content. A simulation study is presented whereby missingness is artificially induced in a Missing At Random (MAR) fashion in order to determine how the effect size of importance measures is affected by various missing value techniques. Importance estimates are derived using MARS, Random Forests and Theil's relative importance.

Retaining Incomplete Data Records for Market Research Estimation: CART Decision Tree Imputation Techniques for Data Sets with Missing Values

✱Ingo Bentrott, University of Technology, Sydney, School of Marketing, Broadway, International 2007 Australia, *Ingo.bentrott@uts.edu.au*

Key Words: Missing Data, CART, Imputation, Marketing, Decision Trees

Survey nonresponse is increasing in both traditional and web based marketing surveys. Missing data will increase variance of the estimates from parametric marketing research analysis. Although parametric missing data recovery methods such as MI and EM can preserve the variance in the data, these techniques work best when the data is missing completely at random (MCAR). The MCAR assumption is rarely met in practice as there is a multivariate nature to most all marketing data sets. The problem with parametric methods of missing data imputation is exacerbated when the data set contains a mixture of data types (i.e., nominal and ratio) and the data space contains a high degree of interactivity. This paper looks at a classification tree nonparametric missing data imputation technique using CART that preserves the natural variance in the data while minimizing the bias of the imputed values.

Structural Equation Model To Diagnose Success Factors for the Commercialization of Information Technology Project

✱So Young Sohn, Yonsei University, Department of Information & Industrial Engineeri, 134 Shinchon-dong, Seoul, 120-749 Republic of Korea, *sohns@yonsei.ac.kr*; Hong Sik Kim, Yonsei University

Key Words: Commercialization, Technology Transfer, Structural Equation Model

In this paper, we propose a structural equation model to identify success factors for commercialization of information and communication technology. We group commercialization scenarios in terms of the type of technology, technology developer and technology receiver. A feedback mechanism is provided to improve the commercialization process for each group based on the observable variables related to the factors. Based on our findings, we suggest some strategies to effectively promote R&D efforts for successful commercialization of each group.

444 From the Golden Ratio to Elephants: A Collection of Viewpoints on Teaching Statistics

Section on Statistical Education

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Golden-Ratio Forms in Piano Sonatas by Mozart and Haydn: A Statistical Analysis

✱Jesper Rydén, Uppsala University, Department of Mathematics, Box 480, Uppsala, International SE-75106 Sweden, *jesper.ryden@math.uu.se*

Key Words: golden ratio, regression, musicology, application

The golden ratio (occasionally called the golden section or the golden mean) has attracted interest from people over the years. Scientists in a wide range of fields have explored or questioned possible presence of golden-ratio proportions. Examples are to be found in nature itself as well as in man-made creations (pieces of art, architecture). If possible, having a sample of observations at hand, a statistical test might be useful for judgment. Some popular beliefs, not to say myths, about the golden ratio can be regarded as misconceptions: statistical tests might be useful for judgment. In this paper we use a regression technique to investigate a problem from musicology: whether the golden-ratio proportion could be said to be present in piano sonatas by W.A. Mozart. The model allows for comparison with works by another master composer, J. Haydn.

Correlation Between the Sample Mean and Variance

✱Ram Shanmugam, Texas State University, 601 University Dr, San Marcos, TX 78666, *rs25@txstate.edu*

Key Words: skewness, kurtosis, normal and non-normal data, count distribution, continuous distribution

This presentation obtains a general formula to find the correlation coefficient between the sample mean and variance. Several particular results for major non-normal distributions are extracted to help students in classroom, clients during statistical consulting service.

Student Attitudes Toward Statistics at Augsburg College

Milo Schield, W.M. Keck Statistical Literacy Project; ✱Cynthia Schield, W. Keck Statistical Literacy Project, *milo@pro-ns.net*

Key Words: statistics education, statistical literacy

Macnaughton argued that the primary goal of an introductory statistics course should be to give students "a lasting appreciation for the value of statistics." Schau argued that student attitudes are "the other important outcome." Schau has designed a reliable instrument for measuring changes in student attitudes on four scales: difficulty, affect, cognitive competence and value. These changes are obtained by comparing the results of a pre and post survey for a given student. Repeated applications of this instrument have found that students see less value in statistics after taking the course than they did before taking the course. This paper presents the results of Schau's longitudinal survey at Augsburg College for students taking traditional statistics and for students taking statistical literacy.

A Study on the Association Between Extraneous Variables and Student Evaluations

✱Lewis Shoemaker, Millersville University of Pennsylvania, Deptartment of Mathematics, Millersville, PA 17551, *lshoemaker@millersville.edu*

Key Words: student evaluations, weighted regression, grades

Student evaluations from 20490 classes over a period of 16.5 years from Spring 1990 through Spring 2006 are analyzed to study the association (if any) between certain extraneous variables and faculty evaluations. The primary methods of analyses used in the study are weighted regression and analysis of covariance. The analysis indicates a moderate relationship between overall evaluation of the instructor and other extraneous variables, including expected grade.

Highlights of Thinking Statistically: Elephants Go to School

✱Sarjinder Singh, The University of Texas at Brownsville and Texas Southmost College, Department of Mathematics, 80 Fort Brown, Brownsville, TX 78520, *sarjinder@yahoo.com*

Key Words: Teaching Statistics, Learning with Fun, K12 level kids alert!, ISBN 07575-3738-7, KENDALL/HUNT PUBLISHING

Carbno (2007), {*Technometrics*, 49(4), pp.496}, wrote, "Sarjinder Singh has produced a delightful introductory statistics text. I had thought that I might find these silly (elephant) examples annoying, but instead I found myself entertained, instructed, and amused. Carbno also added, "Hats off to Singh for taking a daring, innovative approach and doing an absolutely superb job of it." Carbno recommended it for a high school or university course. Finally, he concluded that learning should be fun, and the elephant-book confirmed that it is possible. Kozak (2006), {*Statistics in Transition*, 7(6), 1407-9}, wrote, "If you want to study statistics, if you want to teach statistics, or if you just want to have some fun - you should read this book. You won't regret!" In this talk, I will demonstrate a couple of pages from the book. Some new work, expected to be published soon, will be discussed.

Opportunities with the Journal *Teaching Statistics*

✱Roger W. Johnson, South Dakota School of Mines & Technology, Department of Mathematics & Computer Science, 501 East St. Joseph Street, Rapid City, SD 57701, *Roger.Johnson@sdsmt.edu*

Key Words: Author, Referee, Publication

Opportunities with the Journal *Teaching Statistics* - as author or referee, are presented. *Teaching Statistics* (*http://www.rsscse.org.uk/ts/*), now published by Wiley-Blackwell, has appeared in print three times a year since 1979 with all articles available at *http://www.blackwellsynergy.com/loi/test* (this site will migrate from Synergy to Interscience July 1 2008 as a consequence of the Wiley-Blackwell merger). The first issue of 2008 and the prize-winning C. Oswald George articles are, in fact, are available from this link without charge. The Royal Statistical Society Centre for Statistical Education at Nottingham Trent University serves as home for the journal.

A Review of Poster Presentations Displayed at JSM 2007

✱Edward R. Mansfield, The University of Alabama, Box 870226, Tuscaloosa, AL 35487-0226, *emansfie@cba.ua.edu*

Key Words: Poster presentation, Teaching, Communication

Methods of presentations at JSM have evolved dramatically over many years as presenters strive to speak out and reach out to colleagues. In addition to the growing and effective use of video projection at invited and contributed sessions, poster sessions are becoming more popular. This medium itself is in a rapid evolutionary process as presenters seek better methods of communicating with a transient audience. The 2007 JSM presentation by Di Cook, "Improving Statistical Posters," catapulted the improvement process forward by offering guidelines and illustrations for how use a poster session effectively; view this at *www.amstat.org/meetings/jsm/2008/pdfs/ImprovingStatisticalPosters.pdf*. The current presentation focuses on methods of improving interaction between authors and audience by reviewing and evaluating some of the posters that were displayed at JSM 2007.

445 Variables Selection and Dimension Reduction

Section on Statistical Computing, WNAR
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Principal Component Analysis and Subspace Detection Under Random Projection

✱Qi Ding, Boston University, 111 Cummington Street, Boston, MA 02215, *qiding@math.bu.edu*; Eric Kolaczyk, Boston University

Key Words: Random matrix theory, Principal component analysis, Eigenvalue distribution, Spiked covariance model, Johnson-Lindenstrauss Theorem

Random projection is widely used as a method of dimension reduction. In recent years, its combination with standard techniques of regression and classification has been explored. Here we examine its use with principal component analysis (PCA) and subspace detection methods. Specifically, we show that, under appropriate conditions, with high probability the magnitude of the residuals of a PCA analysis of randomly projected data behaves nearly the same as that of the residuals of a similar PCA analysis of the original data. Our results indicate the feasibility of applying subspace-based anomaly detection algorithms to randomly projected data, when the data have a covariance of an appropriately compressed nature. We illustrate in the context of computer network traffic anomaly detection.

Summarizing X-Ray Stellar Spectra

✱Hyunsook Lee, Harvard-Smithsonian Center for Astrophysics, 60 Garden St. MS-70, Cambridge, MA 02138, *hlee@head.cfa.harvard.edu*; Vinay Kashyap, Harvard-Smithsonian Center for Astrophysics

Key Words: X-ray stellar spectra, dimension reduction, hierarchical clustering, missing data, subspace clustering, algorithms

XAtlas is a spectrum database made with the High Resolution Transmission Grating on the Chandra X-ray Observatory, after painstaking detailed emission measure analysis to extract quantified information. Here, we explore the possibility of summarizing this spectral information into relatively convenient measurable quantities via dimension reduction methods. Principal component analysis, simple component analysis, projection pursuit, independent component analysis, and parallel coordinates are employed to enhance any patterned structures embedded in the high dimensional space. We discuss pros and cons of each dimension reduction method as a part of developing clustering algorithms for XAtlas. The biggest challenge in analyzing XAtlas was handling missing values that are of astrophysical importance. This research was supported by NASA/AISRP grant NNG06GF17G and NASA contract NAS8-39073.

Destructive Effect of the Noise in Principal Component Analysis

✱Katerina Tsakiri, University at Albany, 1400 Washington Avenue, Department of Mathematics and Statistics, Albany, NY 12222, *kt125852@albany.edu*; Igor G. Zurbenko, University at Albany

Key Words: multivariate analysis, principal component, eigenvector

We are considering statistical multivariate analysis of m normally distributed random variables when they were affected by the independent noise. Principal components determine from joint distribution of original sample affected by noise can be essentially different compare with principal component determine from original sample. Examples are provided. Asymptotic properties of covariance matrix affected by noise are investigated when value of the noise is diminishing. When main eigenvalues of original covariance matrix are all distinguishable the effect of small noise proved to be neglectable. Simulation study is provided to determine critical levels of nearly correct determination of principal component in the noisy case. Result is applied to the analysis of the main factor in air pollution data.

Group Variable Selection via Hierarchical LASSO and Its Oracle Property

✱Nengfeng Zhou, The University of Michigan, 4852 Washtenaw Ave. B11, Ann Arbor, MI 48108, *nfzhou@umich.edu*; Ji Zhu, The University of Michigan

Key Words: Group selection, LASSO, Oracle property, Regularization, Variable selection

In many engineering and scientific applications, predictor variables are grouped, for example, in biological applications where assayed genes or proteins can be grouped by biological roles or biological pathways. Existing successful group variable selection methods have the limitation of selecting variables in an "all-in-all-out" fashion (i.e., when one variable in a group is selected) all other variables in the same group are also selected. In this paper, we develop a new group variable selection method that not only removes unimportant groups effectively, but also keeps the flexibility of selecting variables within a group. We also show that the new method offers the potential for achieving the theoretical "oracle" property for both cases where the number of predictors is fixed and the number of groups diverges.

Model Assessment and Selection for Latent Transition Model

✱Hsiu-Ching Chang, Michigan State University, 410 W Saginaw Street, East Lansing, MI 48823, *changh13@msu.edu*; Hwan Chung, Michigan State University

Key Words: Model Selection, Simulation, Latent Transition Model, Longitudinal, Stage-sequential

The Latent Transition Model (LTM) has recently been applied to behavioral and biomedical research to identify individuals' underlying stage-sequential processes. LTM derives from the family of latent class analysis in which manifest items are treated as fallible indicators of unseen states. Specifically, LTM is specified to estimate simultaneously the probability of individuals' membership in latent classes and their transition rate among these classes from occasion to occasion. As a particular example of finite mixtures, model selection is a difficult challenge facing users of LTM: different available methods often suggest different solutions, yielding ambiguity in model selection. Therefore, it is important to identify which method optimally recovers the true model under a variety conditions. In this talk, we review current available methods and compare their performance.

Early Thresholding for High-Dimensional Linear Regression and Variable Selection

✱Xinge J. Jeng, Purdue University, 250 N. University Street, Department of Statistics, West Lafayette, IN 47907-2066, *zheng4@stat.purdue.edu*; Michael Y. Zhu, Purdue University

Key Words: Early Thresholding, Covariance Regularization, Estimation Efficiency, Variable Selection

We introduce a general procedure called Early Thresholding to improve the performance of shrinkage methods such as Lasso and Dantzig selector in high dimensional regression. It is known that in high dimensional regression, singularity of sample covariance matrices can compromise the performance of most estimators. On the other hand, in many applications, a large amount of predictors are weekly correlated. Early Thresholding applies an additional regularization step in estimating the sample covariance matrix, in order to induce a sparse structure. Both theoretical and simulation results show that Early Thresholding can improve estimation and variable selection when combined with Lasso and Dantzig selector.

Shrinkage and Model Selection with Correlated Variables via Weighted Fusion

✱Zhongyin J. Daye, Purdue University, 250 N. University Street, Department of Statistics, West Lafayette, IN 47907-2066, *zdaye@stat.purdue.edu*; Xinge J. Jeng, Purdue University

Key Words: Elastic net, Lasso, Multicollinearity, p>>n problem, Regression, Variable selection

Regression with correlated variables presents a challenging problem in high dimensionality. In this talk, we propose the weighted fusion, a new penalized regression and variable selection method for data with correlated variables. The weighted fusion can potentially incorporate information redundancy among related variables. When the number of predictors p is larger than the number of observations n, weighted fusion also allows the selection of more than n variables in a motivated way. We present grouping effect and consistency results for weighted fusion. Further, we demonstrate real data and simulation examples to show that weighted fusion can improve variable selection and prediction accuracy.

446 Topics in Regression and Dimension Reduction

Section on Physical and Engineering Sciences, Section on Nonparametric Statistics, Section on Quality and Productivity

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

On Dirichlet Regression Residuals

✱Rafiq Hijazi, United Arab Emirates University, CBE, Department of Statistics, B.O.Box 17555, Al-Ain, United Arab Emirates, *rhijazi@uaeu.ac.ae*

Key Words: Compositional data, Dirichlet regression, Deviance residuals

Dirichlet regression, besides logratio regression, has been widely used in modeling compositional data. In this paper, several approaches to calculate residuals for Dirichlet regression are considered. The empirical distributions of the developed residuals are discussed and a Monte Carlo simulation study to evaluate their behavior is performed. Finally, an illustrative example with real data is presented and discussed.

Cluster-Based Lack-of-Fit Test for Nonlinear Regression Models

✱Wijith P. Munasinghe, Kansas State University, 101 Dickens Hall, Kansas State University, Manhattan, KS 66506-0802, *wijith@ksu.edu*; James W. Neill, Kansas State University; Forrest R. Miller, Kansas State University

Key Words: Lack of Fit, Nonlinear Regression, Nonreplication, Maximin Power Clusters

Likelihood ratio tests are used to check a proposed nonlinear regression model for lack of fit. In particular, cluster-based lack-of-fit tests for nonreplicated linear models, as presented by Christensen (1989,1991), have been generalized for nonlinear models with additive error. In addition, cluster selection is based on a maximin power clustering criterion given by Miller, Neill and Sherfey (1998,1999) adapted to the nonlinear case. Simulation results for a two parameter exponential model with one predictor and normal errors indicate that the testing procedure has good power for detecting between-cluster lack-of-fit. Notably, the simulation study shows that for such good power, cluster selection varies according to different points along the expectation surface. The constructed full model uses the convexity of a class of fuzzy clusterings to form a continuous alternative model.

Coresidual Analysis of Linear Models

✱George R. Terrell, Virginia Polytechnic Institute and State University, Statistics Department, Virginia Tech, Blacksburg, VA 24061, *terrell@vt.edu*

Key Words: Regression, ANOVA, Robust methods, Sensitivity analysis

We will show that almost any linear location model (regression, ANOVA, etc.) may be reformulated as an equivalent problem that targets the influence of each observation on the fitting criterion. This alternative formulation corresponds to dual methods in optimization theory, and to shadow prices in economics. Under many circumstances, such as the case of many regressors, the new formulation is computationally less challenging than the original problem. We propose new ways of dealing with issues of design, of robustness, of stability of estimates, and of model selection.

Estimation and Testing of Linear Treatment Effects from Matched Pair Data

✱Cornelis J. Potgieter, University of Johannesburg, PO Box 524, Auckland Park, Johannesburg, International 2006 South Africa, *nelisp@uj.ac.za*; Fred Lombard, University of Johannesburg

Key Words: Matched pairs, Quantile comparison function, Location and scale parameters

In the analysis of matched pairs it is standard practice to reduce the data to pairwise differences and to identify the mean difference with the treatment effect. In this talk a more general measure of treatment effect, the quantile comparison function (QCF) which is essentially a QQ-plot is considered in the context of paired samples (Lombard (2005), *Technometrics*). Specifically, we consider a straight line QCF which is equivalent to a treatment effect consisting of a change of location and/or scale. Permutation tests for the straight-line hypothesis are considered and their powers compared via Monte Carlo simulation. Assuming that the QCF is a straight line, we consider joint estimation of the intercept and slope using sample moments and quantiles and minimum distance methods. Application of the procedures is illustrated in the analysis of parallel coal samples from a common source.

The Use of Singular Value Decomposition and Infrared Spectroscopy To Study Protein Folding Dynamics

✱Taylor Pressler, Mount Holyoke College, South Hadley, MA 01075, *taylor.pressler@gmail.com*

Key Words: Spectroscopy, Protein Misfolding, Singular Value Decomposition, Statistical Modeling, Chemistry, Time Series

Singular Value Decomposition (SVD) is a type of component analysis that can be used to decompose a real m x n matrix. SVD is very useful when analyzing large sets of spectroscopic data collected as a function of time. Infrared Spectroscopy is a technique which is widely used in biophysical chemistry research of the misfolding of proteins, which can be the underlying cause of many diseases. A time series of IR spectra is regarded as a matrix, in which each column corresponds to a spectra taken at a fixed time. A useful way to analyze the data is to decompose the matrix into a sum of terms by SVD. The individual terms in this decomposition then hold a real physical significance. Current research models for SVD and IR are examined and errors are found which ultimately lead to the over-interpretation of physical data.

The Determinant Ratio as a Standardized Index of Data Volume

✱Jason van der Horst, Brigham Young University, *jasonvdh@gmail.com*; Benjamin Heuston, Brigham Young University; Scott Cannon, Brigham Young University; Brian Mead, Brigham Young University; Bruce L. Brown, Brigham Young University

Key Words: determinant, data volume, multicollinearity, sphericity

As pointed out by Rencher (2004), the determinant of a covariance matrix is an index of generalized variance. That is, it quantifies data volume within a multivariate dataset. When standardized, it becomes the determinant ratio, an index of relative data volume adjusted for total variance. Using the assumption of equality of covariance, we examine the relationship between data volume and magnitude of covariance for various matrix sizes from a 4 x 4 matrix to a 20 x 20 matrix. By collectively comparing relative multicollinearity/sphericity across covariance matrices, one can ascertain the maximum number of variables to be used before one experiences diminishing returns, thus allowing researchers to determine at what point they can begin to throw variables out of their model. This will help ensure modeling efficiency by explaining the most variance with the least number of variables.

447 Nonparametric and Semiparametric Regression and Related Problems ●▲

Section on Nonparametric Statistics, IMS

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Spline Confidence Bands for Variance Function via Oracle Efficient Estimation

✱Qiongxia Song, Michigan State University, Dept. Statistics, Wells Hall A410, East Lansing, MI 48824, *song@stt.msu.edu*; Lijian Yang, Michigan State University

Key Words: B spline, heteroscedasticity, knots, nonparametric regression, oracle efficiency, variance function

Asymptotically exact and conservative confidence bands are obtained for possibly heteroscedastic variance function, using piecewise constant and piecewise linear spline estimation, respectively. The variance estimation possesses oracle efficiency and the widths of the confidence bands are of optimal order. Simulation experiments provide strong evidence that corroborates the asymptotic theory while the computing is extremely fast. As illustration of the applicability of the methods, the linear spline band has been applied to test for heteroscedasticity in a fossil data and in the motorcycle data.

Inference on Random Effect Variances with Application to the Assessment of Nonlinearity

✱Sonja Greven, Johns Hopkins University, *sgreven@jhsph.edu*; Ciprian M. Crainiceanu, Johns Hopkins University; Fabian Scheipl, Ludwig-Maximilians-University Munich; Helmut Küchenhoff, Ludwig-Maximilians-University Munich; Annette Peters, Helmholtz Zentrum München - German Research Center for Environmental Health

Key Words: Mixed Model, Random Effect, Penalized Splines, Nonparametric Regression, Non-standard problem

Linear mixed models (LMMs) are a powerful inferential tool. One application are mixed model penalized splines used for nonparametric regression. We are interested in inference on random effect variances. An important special case is testing for polynomial regression against a smooth alternative. Our motivating application is the assessment of nonlinearity for air pollution dose-response functions. Testing for a zero variance is nonstandard due to the tested parameter on the boundary of the parameter space. Furthermore, observations are generally not independent in LMMs. We compare several tests in extensive simulations. We find that our fast finite sample approximation based on the restricted likelihood ratio test statistic outperforms all other tests considered. The class of models covered includes models with moderate numbers of clusters, unbalanced designs, or nonparametric smoothing.

Smooth Regression in Model Selection-Shrinkage and Average

✱Jingyi Liu, University of California, Davis, 4400 Solano Park Cir. Apt 4123, Graduate Group of Biostatistics, Davis, CA 95616, *jyliu@wald.ucdavis.edu*

Key Words: Shrinkage, Average, Model fitting

When a regression problem contains many predictor variables, a regression equation based on a few variables will be more accurate and certainly simpler. Shrinkage methods are widely used in the literature when most information is compressed into the first N dimensions. However, if the regression coefficients show very little difference from each other, it is not easy to choose a few dimensions to represent the data information, especially when the number of predictors has the same order as the number of observations. We develop a method called smooth regression, which includes both shrinkage and average procedure. Simulation results show that smooth regression always outperforms shrinkage alone in terms of Mean Square Error (MSE). Asymptotic approximation of the MSE of the proposed method is derived with certain assumptions of shrinkage and average functions.

AdaSmoothing: Adaptive Smoothing Spline and Its Applications

✱Aijun Zhang, The University of Michigan, 439 West Hall, 1085 South University, Ann Arbor, MI 48109, *ajzhang@umich.edu*

Key Words: AdaSmoothing, nonparametric regression, credit risk, stochastic volatility

Smoothing spline is among the most popular nonparametric regression methods for function estimation, and it has been derived elegantly based on the theory of reproducing kernel Hilbert space. In this paper, we study the adaptive version of smoothing spline by making the smoothing parameter locally adaptive to the nonhomogeneous degree of smoothness at different locations. We show that the shape of such adaptive parameter is determined by the stochastic volatility of a derivative process. As an application, we give an example of using AdaSmoothing in modeling loss-given-default for credit risk portfolio.

Local Polynomial Composite Quantile Regression

✱Bo Kai, The Pennsylvania State University, 326 Thomas Building, University Park, PA 16802, *buk126@psu.edu*; Runze Li, The Pennsylvania State University; Hui Zou, The University of Minnesota

Key Words: nonparametric estimation, local polynomial regression, composite quantile regression, asymptotic efficiency

Nonparametric regression is a useful statistical tool to explore fine features in the data, and has been applied in various disciplines. In this talk, we propose local polynomial composite quantile regression for nonparametric regression models. We derive the asymptotic bias, variance and normality of the proposed estimate. Asymptotic relative efficiency of the proposed estimate to the local polynomial regression under the least squares loss is investigated. We show that the proposed estimate can be much more efficient than the local polynomial regression estimate with the squared loss for various non-normal errors, and is almost as efficient as the LS estimate for normal error. Simulation is conducted to examine the performance of the proposed estimates. The simulation results are consistent with our theoretic findings. A real data example is used to illustrate the proposed procedures.

Design-Adaptive Local Polynomial Estimator for the Errors-in-Variables Problem

✱Aurore Delaigle, University of Bristol, Bristol, BS8 1TW United Kingdom, *Aurore.Delaigle@bris.ac.uk*; Jianqing Fan, Princeton University; Raymond Carroll, Texas A&M University

Key Words: kernel, smoothing, measurement errors, nonparametric regression, bandwidth, deconvolution

Local polynomial estimators are very popular techniques of nonparametric regression estimation and have received great attention in the literature. Their simplest version, the local constant estimator, can be easily extended to the errors-in-variables context by exploiting its similarity with the deconvolution kernel density estimator. The generalization of the higher order versions of the estimator, however, is not straightforward and has remained an open problem for the last 15 years. We propose local polynomial estimators of any order in the errors-in-variables context, discuss their asymptotic properties and illustrate their finite sample performance on numerical data examples.

A Simple Semiparametric Method for Estimating ARCH Models

✱Li Wang, The University of Georgia, 223 Statistics Building, University of Georgia, Athens, GA 30602, *lilywang@uga.edu*; Lijian Yang, Michigan State University; Jianhua Huang, Texas A&M University

Key Words: B-spline, foreign exchange returns, knots, news impact curve, semiparametric regression, volatility

For the past two decades, stochastic volatility models have been extensively investigated in various ways. Among those models, the non/semi parametric (G)ARCH model is particularly interesting due to the great flexibility of the models to be used without restricting a certain shapes of the volatility functions. Linton and Mammen (2005) considered a class of semiparametric ARCH models. They proposed an estimation method based on kernel smoothing and profiled likelihood, which is theoretical in nature with an analysis that relied on a complicated solution of a linear Fredholm integral equations. To make the semiparametric ARCH models more appealing, we developed a simple semiparametric method based on polynomial splines and an fast algorithm to implement the method in practice. The estimation procedure has been illustrated by the daily foreign exchange return data.

448 Bayesian Regression Models and Variable Selection Methods ●▲

Section on Bayesian Statistical Science

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Bayesian Generalized Linear Models with Variable Selection for Various Phenotypes

✱Deukwoo Kwon, National Cancer Institute, 6120 Executive Blvd.,, EPS Rm. 7045, Rockville, MD 20852, *kwonde@mail.nih.gov*; Ruth Pfeiffer, National Cancer Institute; Marina Vannucci, Rice University; Maria T. Landi, National Cancer Institute

Key Words: Bayesian GLMs, variable selection, SELDI-TOF MS, highly correlated predictors

In this paper we propose a Bayesian framework for building prediction models using high dimensional data by combining a Bayesian variable selection method with generalized linear models(GLMs) for various phenotypes. Our strategy is ‘divide and conquer’, that is, we perform variable selection first and then, given selected variables we estimate parameters of interest such as regression coefficients in the prediction models. The rationale of this strategy is that when we have a few true predictors as well as many noisy ones, model uncertainty is a much bigger problem than parameter estimation. We consider binary, ordinal, count, and censored survival outcomes. To illustrate our approach we use data on a proteomic profiling study of melanoma, where serum proteins for each subject were analyzed using SELDI-TOF MS data. We also deal with highly correlated predictors in the variable selection.

Bayesian Analysis of Longitudinal Binary Data Using Markov Regression Models with Skewed Links

Seongho Song, University of Cincinnati; ✱Younshik Chung, Pusan National University, Dept. of Statistics, Pusan National University, 30 JangJeondong, Geumjeong-gu,, Pusan, 609-735 Republic of Korea, *yschung@pusan.ac.kr*; Dipak Dey, University of Connecticut; Chansoo Kim, Kongju National University; Junghoon Jang, Korea Food and Drug Administration

Key Words: Bayesian Markov regression model, Correlated Bernoulli process, Skewed link, Latent variables, Reversible jumps MCMC

In this research, we consider nonhomogeneous Markov regression models of unknown order as a means to assess the duration of autoregressive dependence in longitudinal binary data. We describe a subject’s transition probability evolving over time using logistic regression models for his/her past outcomes and covariates. Our main goal is to develop the appropriate probability model for the correlated Bernoulli Process in the presence of covariate information. In this model, we consider the skewed links for the link function given by Chen, Dey and Shao (1999). Then the model parameters and the order of transitions are estimated using reversible jump Markov chain Monte Carlo (MCMC) approach (Green, 1995; Green and Richardson, 1997).

Bayesian Variable Selection for Logistic Regression with Misclassification

✱Stephanie Powers, Baylor University, 9821 Chapel Rd. #621, Waco, TX 76712, *Stephanie_Powers@baylor.edu*; James Stamey, Baylor University

Key Words: Bayesian Variable Selection, Non-Differential Misclassification, Misclassification, Logistic regression

Response misclassification is a problem in many studies with binary outcomes. In a logistic regression where the binary response is subject to misclassification, we investigate the impact ignoring misclassification has on determining a subset of significant covariates. We then develop a Bayesian approach to select significant covariates while accounting for the misclassification. The procedure is applied to a real data example of a smoking cessation study and extensions to potential pharmaceutical problems are discussed.

Wavelet-Based Bayesian Estimation of Partially Linear Regression Models with Long Memory Errors

✱Kyungduk Ko, Boise State University, 1910 University Dr., Boise, ID 83725, *ko@math.boisestate.edu*; Leming Qu, Boise State University; Marina Vannucci, Rice University

Key Words: Bayesian Inference, Long Memory, MCMC, Partially Linear Regression Model, Wavelet Transforms

This paper proposes a wavelet-based Bayesian estimation method of the model parameters and the nonparametric part of partially linear regression models with long memory errors. We employ discrete wavelet transforms in order to simplify the dense variance-covariance matrix of long memory errors. For a fully Bayesian inference, we adopt a Metropolis algorithm within a Gibbs sampler for the simultaneous estimation of the model parameters and nonparametric function. We evaluate performances on simulated data and then show how the proposed model can be applied to real data, by studying Northern hemisphere temperature data which is a benchmark in long memory literature.

A Bayesian Conceptual Predictive Statistic

✱Andrew Neath, Southern Illinois University Edwardsville, Department of Mathematics and Statistics, Edwardsville, IL 62026-1653, *aneath@siue.edu*; Joseph E. Cavanaugh, The University of Iowa

Key Words: model selection, linear regression

The conceptual predictive statistic, Cp, is widely used as a criterion for model selection in linear regression. Cp serves as an approximately unbiased estimator of a discrepancy, a measure that reflects the disparity between the true model and a fitted candidate model. The “best” model within a candidate class is the one for which the discrepancy is minimum. Model selection based on Cp selects that model for which Cp is minimum as “best.” Strict use of the Cp criterion ignores uncertainty inherent to the model selection problem. In this paper, we present a method for handling this uncertainty using Bayesian methods.

449 Regularized Estimation and Inferential Procedures with Applications to Genomics

Biometrics Section, Section on Nonparametric Statistics

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Boosting Accelerated Failure Time Models for Survival Data with High-Dimensional Covariates

✱Zhu Wang, Yale University, 950 Campbell Avenue, Bldg 35A, Mail code 151B, West Haven, CT 06516, *zhu.wang@yale.edu*; C. Y. Wang, Fred Hutchinson Cancer Research Center

Key Words: Boosting, Accelerated failure time model, Buckley-James estimator, Censored survival data, Variable selection

We consider variable selection for the accelerated failure time models with censored survival data. With high-dimensional covariates, boosting has being successfully applied to regression and classification problems for building accurate predictive models and conducting variable selection simultaneously. However, only limited effort has been made with boosting for the accelerated failure time models when censoring occurs. In this paper, boosting with componentwise linear least squares for model selection is presented for right censored time-to-event data. To accommodate censoring, we consider the mean imputation and Buckley-James method. The proposed methods are evaluated and the Buckley-James method is shown to be comparable with the weighted least squares boosting and weighted LASSO using simulated data. The methods are illustrated with microarray gene expression data DLBCL.

Augmenting the Bootstrap To Analyze High-Dimensional Genomic Data

✱Svitlana Tyekucheva, The Pennsylvania State University, 325 Thomas Bldg, University Park, PA 16802, *sut122@psu.edu*; Francesca Chiaromonte, The Pennsylvania State University

Key Words: high-throughput genomic data, smoothing, bootstrap, resampling

Data produced by high-throughput genomic techniques are often high dimensional and undersampled. This is problematic for many statistical methods, because of overfitting and instability. Techniques that require inversion of covariance matrices (e.g. supervised dimension reduction, the assessment of interdependence structures based on graphical models) are obvious examples, but also techniques that don't require such inversion can behave very poorly when the data is scarce relative to the number of variables involved in the analysis. In this talk, we propose a novel approach, the augmented bootstrap (AB), which can be used to overcome undersampling and improve performance in many statistical analyses. The AB works as a computational "smoother," by simultaneously noising and multiplying data points. In this fashion, it combines the spirit of smoothed bootstrap, bagging, and shrinkage.

A General Framework for Bilevel Variable Selection

✱Patrick Breheny, The University of Iowa, *patrick-breheny@uiowa.edu*; Jian Huang, The University of Iowa

Key Words: Lasso, Penalized Regression, Group lasso, SCAD, Variable Selection, Group bridge

Penalized methods are very useful in sparse regression problems. In many applications, variables have a grouping structure that can be incorporated into the penalty. Here, we introduce a general framework that allows penalties such as ridge, lasso, and SCAD to be combined independently at the group and individual levels, and present algorithms to fit these models. This framework incorporates two existing approaches, group bridge and group lasso. The proposed methods are capable of simultaneous selection at the group and individual variable levels and can be applied to linear and general regression models. Simulations indicate that these group approaches are superior to penalties that ignore grouping structures and provide insight into choosing penalties that best reflect study goals. Finally, we apply these approaches to a genetic association study of age-related macular degeneration.

Dimension Reduction in SNP Data Analysis

✱Adarsh Joshi, The University of Texas MD Anderson Cancer Center, Houston, TX , *ajoshi@mdanderson.org*; Jianhua Hu, The University of Texas MD Anderson Cancer Center; Valen E. Johnson, The University of Texas MD Anderson Cancer Center

Key Words: Single nucleotide polymorphisms, Latent Class Models, Pseudo-Bayes Factors, MCMC, Dimension Reduction, Statistical Genetics

Genomewide SNP data typically consists of hundreds of thousands of SNP markers. Analyzing these genomewide data sets is, therefore, not computationally convenient. Our goal is to reduce a discrete SNP dataset to a more computationally manageable data set without losing relevant information. A Latent Structure Model is applied to segments of SNP data. If the model fits well, then the values of the SNPs within that segment are replaced by the corresponding classification from the latent class model. In this way, we reduce the dimension of the data set. Subsequent analyses can thus be applied to the latent class variables instead of the higher dimensional SNP data.

A Parts-Based Semisupervised Approach for Correlating Large-Scale Biological Data with Censored Survival Data

✱Karthik Devarajan, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA 19111, *karthik.devarajan@fccc.edu*

Key Words: gene expression matrix, high-throughput biology, microarray, nonnegative matrix factorization, semi-supervised learning, censored survival data

Current efforts in cancer research focus on predicting responses of patients by analyzing gene expression patterns and identifying genes that cause specific responses, with a view to personalizing treatment. In the past decade, microarray technology has made it possible to simultaneously measure the expression levels of tens of thousands of genes. Recently, there has been an increased interest in linking such large-scale biological data with censored survival times to predict the survival probability of a patient. In this paper, we propose a semi-supervised approach that combines learning theoretic methods with survival models for censored data. Specifically, we investigate the applicability of nonnegative matrix factorization, a machine learning method for dimensionality reduction, in this context. We illustrate our methods via real-life cancer microarray data as well as simulations.

450 Longitudinal, Clustered, and Correlated Data Analysis

Biometrics Section, Biopharmaceutical Section, Section on Survey Research Methods
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Measuring Impact of Nonignorability in Longitudinal Data with Nonmonotone Nonresponse

✱Hui Xie, University of Illinois at Chicago, IL 60091, *huixie@uic.edu*; Yi Qian, Northwestern University

Key Words: Generalized Least Square, Generalized Linear Mixed Model, Nonignorability, Sensitivity Analysis, Tobit model

The analysis of panel data with nonmonotone nonresponse often relies on the assumption of ignorable missingness. It is important to assess the impact of departure from the ignorability assumption on such an analysis. Nonmonotone nonresponse in longitudinal outcomes, however, can often make such sensitivity analysis infeasible because the likelihood function often involves high-dimensional integrals with respect to missing outcomes. We propose extending a method of local sensitivity analysis to measure the potential effect of nonignorability in panel data with nonmonotone nonresponse. The proposed method overcomes computational difficulty because it completely avoids evaluating the difficult-to-evaluate integrals. We demonstrate the validity and computational advantage of the proposed method with a simulation study. We then illustrate the methodology in the Smoking Trend data set.

A Wilcoxon Rank-Sum Test for Clustered Data

✱Sundar Natarajan, VA New York Harbor Harbor Healthcare System, NY 10010, *sundar.natarajan@va.gov*; Jennifer Friedberg, VA New York Harbor Healthcare System; Stuart R. Lipisitz, Brigham & Women's Hospital

Key Words: Generalized Estimating Equations, cumulative logistic regression model, Score test

With clustered data, such as might arise in a cluster randomized study of 2 treatments for blood pressure, since the subjects within a cluster are not independent, we cannot test for equal treatment effects using a Wilcoxon rank-sum type test. With independent subjects, the Wilcoxon rank-sum test can be shown to be a score test for no covariate effect from a proportional-odds cumulative logistic regression model, with the 'continuous outcome' treated as the ordinal outcome, and the group as a dichotomous covariate. With clustered data, we formulate a similar proportional-odds cumulative logistic regression model (continuous outcome is ordinal and the group is dichotomous). Our Wilcoxon rank-sum test extension to clustered data will be the GEE score test for no group effect under a working independence for this cumulative logistic regression model.

A Two-Stage Approach to Survival Analysis with Time-Dependent Covariates from a Time-Varying Means Model: Application of Time to Initiation of Therapy in HIV-Infected Women vs. Men

✱Amanda A. Allshouse, Health Sciences Center University of Colorado Denver, 4200 East Ninth Avenue C245 HSC, Denver, CO 80262, *Amanda.Allshouse@uchsc.edu*; Samantha MaWhinney, Health Sciences Center University of Colorado Denver ; William Feser, Health Sciences Center University of Colorado Denver ; Amie Meditz, Health Sciences Center University of Colorado Denver; Martin Markowitz, Rockefeller University; Susan Little, The University of California San Diego; Frederick Hecht, The University of California San Francisco; Eric Daar, Harbor UCLA Medical Center; Ann Collier, University of Washington; Joseph Margolick, Johns Hopkins University; Michael Kilby, The University of Alabama; Jean-Pierre Routy, McGill University Health Center; Brian Conway, University of British Columbia; John Kaldor, National Centre for HIV Epidemiology and Clinical Research; Jay Levy, The University of California, San Francisco; Robert Schooley, The University of California, San Diego; D.A. Cooper, National Centre for HIV Epidemiology and Clinical Research; Bruce Walker, Partners AIDS Research Center; Douglas Richman, The University of California, San Diego; Elizabeth Connick, Health Sciences Center University of Colorado Denver

Key Words: Natural B-Splines, longitudinal data, Proportional Hazards model

A mixed-effects, flexible B-Spline model was used in modeling longitudinal and survival data. A longitudinal model was specified with natural B-Splines and then fitted data were incorporated into the Cox proportional hazard model. This method was applied to the Acute Infection and Early Disease Research Program database, which included individuals with acute and early HIV-1 infection from multiple centers, and used to assess a difference in time from study eligibility date to initiation of antiretroviral therapy (ART) in 123 women as compared to 2140 men. Longitudinal data were available for viral load, and CD4 count; number of observations per person ranged from 1 to 28, median=3, mean =5. Estimates of the difference in time to ART by sex from this method were compared to estimates from an extended Cox proportional hazards model with time-varying covariates.

Weighting Method for Binary Longitudinal Data with Incomplete Covariates and Outcomes Incorporating Auxiliary Information

✱Kennedy N. Otwombe, University of the Witwatersrand, School of Statistics and Actuarial Science, Private Bag 3, Johannesburg, 2050 South Africa, *Naviava.Otwombe@wits.ac.za*; Jacky Galpin, University of the Witwatersrand

Key Words: Binary, Longitudinal data, EM algorithm, Multiple Imputation, Weights

A method is developed for the longitudinal missing binary data problem with incomplete data for the outcome and covariate on some subjects, but in which the auxiliary information is always observed. This method applies EM algorithm using the method of weights (Ibrahim, 1990 *JASA* 85:765-9) on an augmented data set derived from the observed data in a regression framework. A missing at random mechanism is assumed. Weights are calculated using the augmented data completing the E-step. Parameter estimates are then determined through the M-step, using a weighted generalized linear mixed model approach that accounts for the correlation structure over time. We assume a simple fixed effects model with a random intercept. Using simulated data, we compare the results from a complete case analysis, multiple imputation and the proposed method and show that the proposed method performs better.

Variance Estimation for Correlated Data

✱Eugenio Andraca-Carrera, The University of North Carolina at Chapel Hill, 3101 McGavran-Greenberg CB#7420, Chapel Hill, NC 27599, *andraca@email.unc.edu*

Key Words: Generalized estimating equation, Robust variance estimator, Correlated data

The robust covariance estimator for generalized estimating equations (GEE) is widely used in the statistics and econometric literature. It is known that use of the robust covariance estimator may lead to inadequate confidence interval coverage when the number of independent clusters is small. Kauermann and Carroll (2001) and Mancl and DeRouen (2001) proposed corrected covariance estimators for GEE with improved small-sample properties. We propose new covariance estimators for generalized estimating equations. The new estimators are comparable to those of Kauermann and Carroll and Mancl and DeRouen in terms of bias and width, but improve upon them in terms of variance in many scenarios. We show in simulations that this reduction in the variance of covariance estimators results in interval coverage that is closer to nominal in small samples.

451 Portfolio Analysis, Exchange Rates, Microstructure, and GARCH Models

Business and Economics Statistics Section
Wednesday, August 6, 2:00 p.m.–3:50 p.m.

Robust Portfolio Selection

✱Michael Schyns, University of Liege, HEC-Management School, Bd du Rectorat 7 (B31), LIEGE, International B4000 Belgium, *M.Schyns@ulg.ac.be*

Key Words: Portfolio selection, Min Cov Determinant, Convex optimization, Robust statistics

In many financial problems, small variations in some inputs may result in big changes in the outputs. In this talk, we consider the problem of portfolio selection as suggested by Markowitz. This model relies on a covariance matrix usually estimated using historical returns of the assets under consideration. Gross error in these returns or atypical events occurring in the past could lead to different portfolios with quite different expected returns. Defining methods that do not depend too much on these atypical data is the aim of robust statistics. We will show that some techniques developed in that field are worth applying in our context. More precisely, the covariance matrix of historical data will be estimated with the Minimum Covariance Determinant estimator, computed with a 'smooth' algorithm. This robust Markowitz methodology will be illustrated on real financial data.

Portfolio Choice in Retirement: Health Risk and the Demand for Annuities, Housing, and Risky Assets

✱Motohiro Yogo, University of Pennsylvania, Finance Department, 3620 Locust Walk, Philadelphia, PA 19104-6367, *yogo@wharton.upenn.edu*

Key Words: Aging, Annuity, Health, Medical expenditure, Portfolio choice

This paper develops a consumption and portfolio-choice model of a retiree who allocates wealth among four assets: a riskless bond, a risky asset, a real annuity, and housing. Unlike previous studies that treat health expenditures as exogenous negative income shocks, this paper builds on the Grossman model to endogenize health expenditures as investments in health. I calibrate the model to explain the joint evolution of health status and the composition of wealth for retirees, aged 65 to 96, in the Health and Retirement Study. I use the calibrated model to assess the welfare gains of an actuarially fair annuity market. The welfare gain is less than 1% of wealth for the median-health retiree at age 65, and the welfare gain is about 10% of wealth for the healthiest.

Applications of Statistics in Finance Using the Statistics Online Computational Resource (SOCR)

✱Nicolas Christou, University of California, Los Angeles, 8125 Math Sciences Bldg., Los Angeles, CA 90095, *nchristo@stat.ucla.edu*; Ivo D. Dinov, University of California, Los Angeles

Key Words: efficient frontier, covariance, portfolio risk and return, stock market, Java applets, SOCR

In this paper we present an application of statistics in finance using SOCR. Most, if not all, college students have some familiarity with the stock market and therefore can understand the problem easily. It is the real data analysis that students find interesting and sometimes challenging. Here we explore the building of efficient portfolios using real stock market data through optimization examples of two, or multiple stocks, and how covariance and correlation can help the investor to diversify his or her risk. We discuss why diversification works, but also the problems that arise in portfolio management. Stock market data can be incorporated at any level of Mathematics, Statistics and Economics education. With the use of these newly designed SOCR applets and learning materials, we can enhance the teaching of statistics and portfolio management in varieties of curricula.

A Study on Dependent Relation of Currency Exchanging Rates Between Japanese Yen and Taiwan Dollar

✱Yi-Kuan Jong, St. John's University, Taipei, International Taiwan, *ykjong@mail.sju.edu.tw*

Key Words: VaR, Currency Exchange, Copula

The concept of Value at Risk(VaR) appeared in 1994, it is used to evaluate financial risk. After the New Basel Capital Accord was published, the risk management was emphasized for Bank regulations. It become an important measurement for evaluate financial risk. For oriental currencies, Japanese Yen is the most important one. Its exchange rate to U.S. dollar will influence the other exchange rate of currencies to U.S. dollars. In this study, we want to copula method to model multivariate distribution of the returns of exchange rates for JPN-USD and TWD-USD. Then use Value at Risk to evaluate the potential finance risks we are going to encounter.

Market Depth in Agricultural Futures Markets

✱Julieta Frank, University of Illinois at Urbana-Champaign, *julietafrank@yahoo.com*; Philip Garcia, University of Illinois at Urbana-Champaign

Key Words: market depth, agricultural futures markets, market microstructure

Market depth is the order flow required to move prices by one unit. A natural approach to measure market depth would be to estimate the slope in the price-volume relationship. However, in agricultural futures markets this is not straightforward because buy and sell orders occur in an open outcry pit and are not recorded. Estimation of these unobserved data conditional on transaction prices and volume is performed using a Bayesian MCMC algorithm, the Gibbs sampler, built on Roll's market microstructure model. The degree of market depth in agricultural futures markets has not been investigated to any extent; however research in these markets seems warranted as findings may have a great impact in operational decisions. We provide market reaction curves for hogs and cattle futures contracts trading in the Chicago Mercantile Exchange.

The Information Content of Trades: A Class of Market Microstructure Models

✱Anna Valeva, Western Illinois University, Information Systems and Decision Sciences, 1 University Circle, Macomb, IL 61455, *AK-Valeva@wiu.edu*

Key Words: market microstructure, asymmetric information, trading volume, bid-ask spread, hidden Markov model

We focus on a class of models in which the market specialist exploits the information content of trades to set the bid-ask spread for a given asset in an asymmetric information environment. While the idea for such type of models dates back to the mid nineteen eighties, we adopt an infinite horizon and introduce a dynamic way of quantifying the private information through privileged knowledge about the dividend process associated with the asset. Thus, we have a ''perpetual'' model for the dynamics of the bid-ask spread, in which the market specialist updates her/his beliefs about the unobservable components of the model and tries to infer the privileged information from the volume of trade. We also examine the serial dependence patterns of the bid-ask spread and trading volume, which can be linked to empirical properties of market data.

Orthogonal GARCH Models with Structure Breaks

✱Lingyu Zheng, Temple University, PA 19122, *lzheng@temple.edu*

Key Words: Orthogonal GARCH, Principal Component

Since the seminal paper of Engle (1982), ARCH or GARCH models are now commonly used to describe and forecast changes in the volatility of financial time series. Multivariate extensions of ARCH and GARCH models (MGARCH) are popular in the study of the relations between the volatilities and co-volatilities of several markets. Alexander (2001) illustrates the use of the orthogonal GARCH model, by decomposing the data to a small number of uncorrelated principal components which follow univariate GARCH models. The model is very successful in reducing the number of parameters, but it ignores temporal structure breaks. We improved the orthogonal GARCH model by allowing structure breaks with known or unknown breaking points. We applied the model to empirical data and proved its reliability.

452 Statistical Issues in Medical Research ●

Biopharmaceutical Section, WNAR, Section on Health Policy Statistics, Biometrics Section

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

'Promotion Bias' in Clinical Research Design, Analysis, and Reporting: A Challenge to Statistical Science

✱Peter B. Imrey, The Cleveland Clinic and Case Western Reserve University, 9500 Euclid Avenue/Wb4, Cleveland, OH 44195, *imreyp@ccf.org*

Key Words: clinical trials, clinical trial registries, medical product regulation, statistical ethics, ghost authorship, seeding trials

Statistical design and analysis are central to regulation and dissemination of information about medical products. An extensive literature now documents concerns as to the integrity of clinical research on some products, and suggests that medical information flow is notably distorted by researcher conflicts-of-interest and practices that subordinate science to marketing. These include programs of product seeding trials, ghost authorships of research and review papers for opinion leaders, ghost management of research by communications firms, and selective publication of results of clinical studies. While their extent is unclear, the existence of these abuses is uncontested. This paper argues that unlabeled commingling of science and advocacy in medical research deeply threatens the statistical profession, and should be more strongly and publicly discouraged and resisted by statisticians.

Statistical Considerations for Comparability Studies

✱Frank Ye, Amgen, Inc, 1 Amgen Ct, MS 92-2-B, Thousand Oaks, CA 91360, *fye@amgen.com*; Steve Mitchell, Amgen, Inc

Key Words: Biopharmaceutical, Comparability, CMC, Process Change, Equivalency

Comparability is a process of assuring continuous safety and efficacy by demonstrating that previous product results remain applicable as manufacturing process, product formulation and/or analytical testing method changes are made. Comparability studies can be designed to evaluate performance parameters that are nonprofile (e.g., lot release data) or profile (e.g., stability data) in nature. Comparability studies can also be grouped into two statistical approaches: an approach that requires an explicit and meaningful comparison of distributions or an approach that is based on pre-defined acceptance criteria that reflect the historical data trend under the old process. Statistical techniques such as tolerance intervals and equivalency tests will be discussed and recommended as computational tools for comparability studies.

The Use of Control Charts for the Prospective Monitoring of Long-Term Vaccine Efficacy

✱J. Brooke Marshall, Merck Research Laboratories, M/S UG1CD-44, 351 N. Sumneytown Pike, North Wales, PA 19454, *brooke_marshall@merck.com*; David Radley, Merck Research Laboratories; Lisa C. Lupinacci, Merck Research Laboratories

Key Words: public health surveillance, statistical process control, disease monitoring, cumulative sum, vaccine effectiveness

Once a vaccine has been licensed for widespread use to reduce the incidence of a specific disease, it is important to monitor the breakthrough cases of this disease in the vaccinated population to determine if and when vaccine efficacy wanes, and when a booster dose may be advisable. Several control charting methods can be used to monitor breakthrough cases of disease over time, including Shewhart-type and Cumulative Sum charts. The use of these methods will be discussed in the context of a study being conducted to monitor the long-term efficacy of a new vaccine. Using control charts in this type of application presents a unique challenge because subjects are vaccinated at different times, and it is important to consider their time since vaccination in these methods. A scheme for modifying the standard implementation of the control charts to account for this will also be discussed.

Meta-Regression of Placebo Response

✱Dawn Odom, RTI Health Solutions, PO Box 12194, Research Triangle Park, NC 27709, *dodom@rti.org*; Jianmin Wang, RTI Health Solutions; William Irish, RTI Health Solutions

Key Words: meta-analysis

In traditional meta-analysis and by extension meta-regression, treatment effect sizes (e.g., odds ratio, relative risk) are often estimated using a model with random effects to account for heterogeneity between studies. In a direct comparison, meta-regression typically examines the relative effect of treatment versus placebo while controlling for study-level or treatment-level predictors. However, it may be of interest to estimate the expected placebo response to establish a more reliable reference value. This study will demonstrate our modifications to the standard meta-regression approach and present findings based on a literature review of clinical trials.

Acceptance Criteria in the Evaluation of Blood Products

✱Tie-Hua Ng, U.S. Food and Drug Administration, 1401 Rockville Pike, 200S, HFM-219, Rockville, MD 20852-1448, *tiehua.ng@fda.hhs.gov*

Key Words: acceptance criteria, specifications, blood products

Biologics License Application (BLA) for biologics products are submitted to the Center for Biologics Evaluation and Research (CBER) of the US Food and Drug Administration (FDA) for review. Office of Blood Research and Review (OBRR) within CBER review blood and blood related products. An overview of blood and blood related products will be presented. Determination of acceptance criteria in the evaluation of blood and blood related products will be discussed.

Roles of Meta-Analysis in Global Drug Development

✱Peter H. Zhang, Otsuka Pharmaceutical Development & Commercialization, Inc., 2440 Research Boulevard, Rockville, MD 20850, *peter.zhang@otsuka.com*; Rose Hu, Fast-Track Drugs and Biologics

Key Words: Meta-Analysis, Adaptive Design, Clinical Development, Global Drug Development, Random-effect, medical Research

The impact of meta-analysis has grown enormously not only in medical research but also in global drug development as the need for medical research and clinical practice to be based on the totality of relevant and sound evidence has been increasingly recognized. In global drug development, however, meta-analysis are limited in accepted regulatory roles to provide a more precise estimate of the overall treatment effects, to evaluate whether overall positive results are also seen in pre-specified subgroups of patients, to evaluate safety in a subgroup of patients, or a rare adverse event in patients, and to evaluate apparently conflicting study results. In this paper, we review highlights of recent development in meta-analysis in global drug development to provide confirmatory evidence in both efficacy and safety in order to get regulatory approval and improvement labeling.

453 Contributed Poster Presentations

Biometrics Section, Biopharmaceutical Section, Section on Bayesian Statistical Science, Section on Quality and Productivity, Section on Risk Analysis, Section on Statistics and the Environment, Section on Statistics in Sports

Wednesday, August 6, 2:00 p.m.–3:50 p.m.

A Model of Playing Time for Pitchers in Major League Baseball

✱Michael Greene, Deloitte Consulting, LLP, 70 Burbank St Apt 18, Boston, MA 02115, *mike@quantsportsguy.com*

Key Words: Baseball injuries, playing time model, prediction of playing time

Injuries to baseball players heavily influence the statistical output of a baseball player over time. Missing time due to injuries complicates the prediction of a player's future performance. Collecting comprehensive data on injuries is difficult, thus no complete databases are available. Injured players often continue to play, which can exacerbate injuries and reduce performance. We propose a model for expected playing time for pitchers based on performance and player characteristics. The model is constructed using data from the 2005-2007 seasons and blind validation is applied to demonstrate the power of the model. This flexible model allows for the calculation of expected playing time for pitcher on a weekly basis during the season. Such a tool is useful not only for analysis, but also for teams to monitor players for possible injuries based on decreases in performance.

Sample Size Estimation for Equivalence and Noninferiority Tests

✱Jie Wang, Baylor University, 10661 Steppington Dr. #3100, Dallas, TX 75230, *jie_wang@baylor.edu*

Key Words: equivalence test, non-inferiority test, sample size estimation, Bayesian

Sample size estimation for the equivalence and noninferiority tests for diagnostic methods is considered in Bayesian approach. The goal of the study is to test whether an experimental treatment of interest is equivalent to, or not inferior to another reference treatment, which may or may not be a gold standard. Sample sizes are chosen by the model performance criteria of average posterior variances and lengths. In the absence of gold standard, sample sizes are evaluated by the ratio of marginal probabilities of two treatments with a noninformative Beta prior; whereas, in the presence of gold standard, sample sizes are evaluated by the measures of sensitivity and specificity, with a non-informative Dirichlet prior. The methodology and results for the two-arm trials can be further extended to three-arm trials by comparing two experimental treatments against a gold standard.

Bayesian Inference on Relative Risk Incorporating Multiple Covariates

Qianqiu Li, Wyeth Research; ✱Yonggang Zhao, Wyeth Research, *yonggangzhao69@yahoo.com*

Key Words: Relative risk, Bayesian inference, Poisson regression, MCMC

Relative risk, defined as ratio of two incidence rates, is a common quantity for comparing the rates in two populations. In the presence of confounding factors, Poisson regression is usually employed to adjust relative risk. This paper implements a Bayesian Gamma-Poisson regression model to infer adjusted relative risk. Computations are carried out using Markov Chain Monte Carlo (MCMC) methods. Comparisons are made with the frequentist approaches: lognormal-Poisson regression, Log-Binomial regression, negative-Binomial regression, as well as stratified relative risk.

Bayesian Data Assimilation for Parameter Estimation in Hydrological Systems

✱Darl D. Flake, II, Utah State University, Department of Mathematics and Statistics, 3900 Old Main Hill, Logan, UT 84322, *darl.flake@aggiemail.usu.edu*; Mevin B. Hooten, Utah State University; Luis A. Bastidas, Utah State University

Key Words: data assimilation, Bayesian statistics, hydrology

Conventional approaches to parameter estimation for hydrological systems often involve mechanistic model calibration. Because of the potential complexity of these models, Bayesian methods provide an appropriate framework to assimilate data into the calibration process and thereby allow us to formally account for uncertainty while estimating parameter distributions. We compare and contrast the effectiveness of two such runoff models for appropriately characterizing the physical system under study (i.e., HyMod and the Sacramento Model). Specifically, we present a model formulation that explicitly accommodates the non-negative support of the data and focus the application on historical data from the Leaf River basin in Mississippi. We formally evaluate the utility of the statistical model as well as the mechanistic models involved and discuss additional aspects of the project.

Bayesian Nonparametric Polya Tree Mixture Models with Application to Random Effects Meta-Analysis

✱Adam Branscum, University of Kentucky, College of Public Health, 121 Washington Ave, Lexington, KY 40536, *abran3@email.uky.edu*; Timothy Hanson, The University of Minnesota

Key Words: Bayesian nonparametrics, Polya trees

A goal in meta-analysis is estimation of a single effect measure using data from several studies that are each designed to address the same scientific inquiry. Because studies are typically conducted in geographically disperse locations, the statistical analysis of meta-analytic data often involves the use of random effects models that account for study-to-study variability. To eliminate the influence of overly restrictive parametric models on inferences, we develop a novel hierarchical Bayesian nonparametric Polya tree mixture model. We present methodology for testing the Polya tree mixture versus a normal model. An application involving meta-analysis of epidemiologic studies designed to characterize the association between alcohol consumption and breast cancer is presented, which together with results from simulated data highlight the performance of Polya tree mixtures.

A Bayesian Bivariate Random-Effects Meta-Analysis for Two Correlated Outcomes Using Individual Patient Data

Ying Yang, Bristol-Myers Squibb Company; ✱Kao-Tai Tsai, BMS, 34 Baldwin Dr., Berkeley Heights, NJ 07922, *tsai0123@yahoo.com*

Key Words: multiple outcomes, Bayesian, bivariate meta-analysis

Multiple outcomes are often of interest in psychological clinical trials. For example, the Hamilton Depression Rating Scale (HAMD) and the Montgomery and Asberg Depression Rating Scale (MADRS) are two important outcomes. A separate univariate analysis is often used to evaluate treatment effect for each outcome independently. This approach ignores the correlation between outcomes. To fully assess treatment effect and understand the association between treatments and outcomes, we propose an approach which is a single Bayesian bivariate meta-analysis model that can jointly synthesize those endpoints and utilize their correlation. The individual patient data from five double-blinded, placebo-controlled trials using selegiline transdermal system for major depressive disorder is used to demonstrate the method.

Assessing Local Model Fit in Bayesian Regression Models Using the Partitioned Deviance Information Criterion

✱David Wheeler, Emory University, 1518 Clifton Rd, NE Third Floor, Rollins School of Public Health, Atlanta, GA 303022, *dcwheel@sph.emory.edu*; Lance Waller, Emory University; DeMarc A. Hickson, University of Mississippi Medical Center

Key Words: Bayesian statistics, spatial statistics, DIC, diagnostics, HIV

There has been a recent emphasis in the applied spatial statistical literature on models with local covariate effects, instead of more traditional models that represent relationships with static effects across a study area. Bayesian spatially varying coefficient models can be computationally demanding to fit, and diagnostic tools that justify the additional computation effort are welcome. We use a partitioning of the deviance information criterion (DIC) to assess local model fit and observation influence in a Bayesian framework. We map local DIC, deviance residual, and leverage values to assist in model selection, moving from a global regression model to a spatially varying coefficient regression model. We demonstrate the utility of the local diagnostics with an example of HIV prevalence among pregnant women in the Butare province of Rwanda during 1989–1993.

Bayesian Objective Testing of Hardy-Weinberg Equilibrium

✱Brenda Betancourt, University of Puerto Rico, Av Universidad # 4 Torre Norte, San Juan, PR 00925 Puerto Rico, *b.betancourt@uprrp.edu*; Maria E. Perez, University of Puerto Rico

Key Words: Hardy-Weinberg equilibrium, Intrinsic Priors, Hypothesis Testing

Assessment of Hardy-Weinberg equilibrium is one of the basic problems in population genetics and it is far for being closed from the statistical point of view, as recent efforts in this direction prove. The selection of prior distributions for testing Hardy-Weinberg equilibrium is a difficult issue, as we are dealing with a low-dimensional null hypothesis. Recent advances in Objective Bayesian Analysis allow the construction of priors specially suited for Bayesian testing (intrinsic priors). In this work, an intrinsic prior for testing Hardy-Weinberg equilibrium is calculated from the uniform prior using training samples as in Casella and Moreno (2006). The prior obtained is a mixture of Dirichlet distributions. An analysis of sensitivity to different sizes of the training samples is shown.

A Sequential Bayesian Approach to Distributed Source Localization in Wireless Sensor Networks

✱Soma Dhavala, Texas A&M University, *soma@stat.tamu.edu*; Rajesh Talluri, Texas A&M University; Aleksandar Dogandzic, Iowa State University

Key Words: sequential bayesian analysis, data fusion, distributed estimation, wireless sensor networks, source localization

We propose a sequential Bayesian approach to distributed source localization in wireless sensor networks using the Received Signal Strength (RSS) measurements that follow an energy-decay model. The sequential approach to the problem allows the processed measurements to be communicated across the sensors in a distributed fashion that offers flexibility in managing the communication cost. We consider several probabilistic fusion schemes, namely, the Independent Likelihood Pool, the Independent Opinion Pool and the Linear Opinion Pool. We also consider some numerical approximations that can reduce the communication cost. We evaluate the performance of the proposed methods and compare them with existing distributed algorithms such as incremental non-linear least-squares approach and show that our methods outperform them in terms of the MSE and communication cost.

It Really Is Getting Warmer

✱Jun Lu, Texas Tech University, Department of Mathematics and Statistics, Lubbock, TX 79409, *june.lu@ttu.edu*; Anton Kliewer, Texas Tech University; Clyde Martin, Texas Tech University

Key Words: Climate change, Time Series Data, Loess Smoother, Extreme Value Theory

Global warming is gaining attention as a major threat. This paper focuses on weather data collected for one city, Lubbock, Texas, over a time period of 93 years. New temperature records are being set every year in Lubbock and our goal was to determine if these records are typical or a byproduct of a warming climate. To answer these question two different methods were employed to investigate the available data. The first was to fit the data to the Generalized Extreme Value distribution. This method produced inconclusive results but did give the impression that warming exists. To check this belief, a second method involved a time series data processing technique known as "Loess" Smoother to determine the magnitude of the increasing daily temperatures over time. The results suggest a three degree temperature increase in Lubbock over the last 93 years.

Are We Getting More Rain?

✱Anton Kliewer, Texas Tech University, Department of Mathematics and Statistics, Lubbock, TX 79409, *anton.kliewer@gmail.com*; Clyde Martin, Texas Tech University

Key Words: Precipitation, GEV Distribution, Climate change

This paper studies measurable responses to climate change at the local level. Our goal was to determine if extreme weather events are increasing. We quantify these extreme events over a long period of time. It has been predicted that there will be an increase in variation in precipitation in southwest. Large fluctuations exist from year to year for the daily precipitation levels for Lubbock, Texas over the last 93 years. The precipitation data was fit to the Generalized Extreme Value distribution. Numerous computer simulations were conducted to determine the behavior of the GEV distribution given both static and changing data sets. Comparing these results with the weather data from Lubbock, Texas, the conclusion is that while the precipitation levels vary from year to year, there is no long term change, rather there exists a cyclic pattern.

Comparison of Soil Temperatures at Agricola and Pascagoula, MS

✱Madhuri S. Mulekar, University of South Alabama, Department of Mathematics & Statistics, 307 Univ Blvd, ILB 325, Mobile, AL 36688, *mmulekar@jaguar1.usouthal.edu*; Sytske Kimball, University of South Alabama; Jacob V. Sowell, University of South Alabama

Key Words: Soil temperatures, Diurnal temperature range, DTR

Environmental scientists use soil temperatures in various applications, including input to global circulation models and to verify and grountruth satellite observations. Agricultural scientists use them to predict changes in crop and woodland productivity. Soil temperatures fluctuate annually and daily and are affected mainly by variations in air temperature and solar radiation. Data collected by University of South Alabama Mesonet stations located in rural Mississippi over a period of about 600 days was analyzed. These stations recorded soil temperatures at the ground surface and below ground at 5, 10, 20, 50, and 100 cm on the Celsius scale once every minute. Significant seasonal and yearly cycles were observed in soil temperatures. The diurnal temperature ranges showed significant variation by season and location, coastal versus inland; with higher consistency deeper into the soil.

Generalized Linear Modeling Approach to Stochastic Weather Generators

✱Eva M. Furrer, National Center for Atmospheric Research, 1850 Table Mesa Dr, Boulder, CO 80305, *eva@ucar.edu*; Richard W. Katz, National Center for Atmospheric Research

Key Words: GLM, ENSO, Climate, Weather, Precipitation modeling, Downscaling

Stochastic weather generators are a popular method to produce synthetic sequences of daily weather. We demonstrate how an extension of the Richardson model based on generalized linear models (GLMs) can provide a general modeling framework, allowing for the straightforward incorporation of annual cycles as well as other covariates (e.g. an index of the El Nino-Southern Oscillation (ENSO) phenomenon) into stochastic weather generators (downscaling). We apply the GLM technique to daily time series of weather variables (i.e. precipitation as well as minimum and maximum temperature) at Pergamino, Argentina. Besides annual cycles, the fit is significantly improved by permitting both transition probabilities of the first-order Markov chain for daily precipitation occurrence, as well as the means of both daily minimum and maximum temperature, to depend on the ENSO state.

Association Between Fine Indoor Particulate Pollution and Pulmonary Function

✱Robert L. James, Rho, Inc., 6330 Quadrangle Dr., Ste. 500, Chapel Hill, NC 27517, *rjames@rhoworld.com*; Agustin Calatroni, Rho, Inc.; Herman Mitchell, Rho, Inc.

Key Words: Pulmonary function, FEV1, PEF, Particulate pollution, PM2.5

Fine particulate pollution has recently been associated with increased mortality and respiratory diseases. The NIH Inner-City Air Pollution Study is a prospective study designed to investigate the impact of fine particle pollution (PM2.5, <2.5 microns) on lung function measured by FEV1 (forced expiratory volume in one second) and peak expiratory flow (PEF). The study design had several challenges: (1) the design was highly hierarchical, with children measured for fourteen consecutive days during the study's 12, 18, and 24 month visits; (2) the device measuring PM2.5 in each child's house outputs a measure every ten minutes, creating massive amounts of data; (3) two different devices were used to make the pollution measures; and (4) lung function maneuvers are effort dependent and missing data can be related to pollution levels. Each of these topics is illustrated in detail.

Estimating Limit of Detection in Bioassay

✱Changjian Jiang, Monsanto Company, *changjian.jiang@monsanto.com*

Key Words: Limit of Detection, Bio-assay, Response curve, Over dispersion

In environmental samples, bioassays are usually used to evaluate the potency of an agency contained in the sample on living organisms, such as insects or rats. Similar to the bioanalytical methods, the limit of detection is a measure of the sensitivity of the method. Bioassay uses a standard curve over a series of concentrations for inversely predicting the concentration level in the test sample. Similar to bioanalytical assay, the results are usually affected by some additional random factors which in this case are due to the characteristics of the living organism, such as variation among batches or the fit of the response curve over the concentration levels. In this study, we examine the effect that these additional random factors have on the estimation of the limit of detection.

An International Comparison of Drinking Water Regulations

✱Eric Guttorp, University of Washington, 411056 McMahon Hall, Seattle, WA 98185, *eguttorp@u.washington.edu*; Peter Guttorp, University of Washington; Duru Altug, University of Washington

Key Words: Drinking water, Tap water, Environmental regulations

Recently, water has been a growing concern for many people all over the world. What's in our water? Is it safe to drink the water? Is there lead in the water? Some blame this concern on the marketing attempts of international companies introducing bottled water to the public, claiming it to be cleaner than tap water. In reaction to the heightened concerns of the public, many governments consider it their duty to make sure that the public gets quality drinking water.By comparing the data found on Washington State and the United States, Scotland, Sweden, and Turkey, we hope to find a link between government intervention and water quality levels in order to inform the public of how affective governments can be on this issue.

An Adaptive Sampling Procedure for Estimating Heavy Metal Pollutants of Ground Water from Soil

✱Ismaila A. Adeleke, University of Lagos, Department of Actuarial Science, Faculty of Business Administration, Lagos, 9610001 Nigeria, *adeleke22000@yahoo.ca*; Ray Okafor, University of Lagos; Ebeneezer O. Esan, University of Lagos; Kehinde Olayinka, University of Lagos; Aderonke Oyeyiola, University of Lagos; Anthanasius Opara, University of Laos

Key Words: Adaptive Sampling, Heavy Metal Pollutants, Design-Assisted-cum-Model-Based Estimation Approach

We consider an adaptive cluster sampling procedure for assessment of contaminant risk to ground water from soil at Owode-Onirin, Lagos, Nigeria. This method provides a more efficient way of sampling sparse but highly clustered population, and takes the advantage of spatial clustered population to yield more efficient estimates of the concentrations of heavy metals pollutants of ground water from soil. A model-assisted estimation approach that incorporates models into the design-based estimation frame work was considered.

A Bootstrap Model for Estimating the Concentration of Some Heavy Metals To Determine Contaminant Risk to Ground Water from Soil

Ismaila A. Adeleke, University of Lagos; ✱Ray Okafor, University of Lagos, Department of Mathematics, Faculty of Science, Lagos, 9600001 Nigeria, *okaforray@yahoo.com*; Ebeneezer O. Esan, University of Lagos; Kehinde Olayinka, University of Lagos; Aderonke Oyeyiola, University of Lagos; Shakirudeen Odunuga, University of Lagos; Dallah Hamadu, University of Lagos

Key Words: Adaptive Sampling, Heavy Metal Pollutants, Bootstrap Estimation Model, Horvitz-Thompson and Hansen-Horvitz-Type Estimation

Estimation of the concentration of some heavy metals in soil is very important in risk assessment study to determine contaminant risk to water from soil. An adaptive cluster sampling technique was employed and a model-assisted approach that incorporates models into design-based estimation was used. In this paper, a bootstrap procedure is proposed to approximate the distribution of the concentration of heavy metals in the soil. The proposed method overcomes the problems associated with methods requiring distributional assumptions the validity of which are often times extremely difficult to check. We conclude with a comparison of this procedure to the well known Horvitz-Thompson and the Hansen-Horwitz type estimators.

Optimization of Nonlinear Parameters Used To Model Rice Drying

✱Andy Mauromoustakos, University of Arkansas, 101 AGRX, Agri Stat Lab, Fayetteville,, AR 72701, *amauro@uark.edu*; Terry Siebenmorgen, University of Arkansas; Vir Gayanilo, University of Arkansas; Rasty Bautista, University of Arkansas

Key Words: nonlinear modeling, Page Equation, Optimization

Rice drying is an energy-intensive, time-critical, and quality-sensitive operation. Thus, rapid drying with the least cost, without incurring quality losses, is of great interest to the rice industry. High temperature fluidized bed drying offers fast and uniform drying of granular products such as rice. This study was conducted to use nonlinear regression modeling to quantify rice drying rates using air temperatures (T) ranging from (60-90 degrees C) in combination with air relative humidities (RH) ranging from (7-75%), using a specialized drying chamber. The study determines the rates and the rice equilibrium moisture content (EMC), and all other important nonlinear parameters using Page's nonlinear model and subsequently explores the experimental region for the ideal inputs that would provide the Rice industry with the desired T and RH settings.

Challenges and Solutions for Assessing the Long-Term (1989–2007) Effects of a Marine Oil Terminal in Alaska on the Marine Sediment Environment

✱Arny L. Blanchard, University of Alaska Fairbanks, P.O. Box 757220, Institute of Marine Science, Fairbanks, AK 99775-7720, *arnyb@ims.uaf.edu*

Key Words: Marine Ecology, Spatio-temporal Modeling, Environmental Assessment

Long-term assessment (19 years) of marine, sediment-dwelling organisms documents small effects from chronic exposure to hydrocarbons in discharges of treated ballast waters at a marine oil terminal in Alaska. Challenges for this project include detecting small, chronic effects on benthic organisms from hydrocarbons and communicating statistical results to nonscientists. Chronic effects from the low sediment hydrocarbon concentrations include abundance decreases for a small number of sensitive organisms. Analytical methods common in community ecology (multivariate ordination) have failed to adequately demonstrate the small changes. Graphic presentations of statistical results from spatio-temporal modeling and ANOVA clearly demonstrate the small, chronic effects while effectively communicating ecologically relevant trends to the broad audience of this study.

Density Function Solutions of Population Size Models

✱James Matis, Texas A&M University, department of statistics, college station, TX 77843-3143, *matis@stat.tamu.edu*; Thomas Kiffe, Texas A&M University; Timothy Matis, Texas A&M University; Wopke van der Werf, Wageningen University; Alejandro Costamagna, University Minnesota; G. Jerry Michels, Jr., Texas A&M University

Key Words: mechanistic models, aphids, Normal curve, logistic curve, differential equations

We develop two mechanistic models for describing insect population growth curves. The first is a diff. equation in which the rate of population growth decreases linearly over time, reflecting the aging of the host plants, a 'bottom-up' model. The surprising solution, readily obtained using calculus, is a scaled Normal pdf, hence the Normal curve is fitted directly to growth curve data. The second model is a diff. equation in which the death rate is a function of the cumulative past population size, reflecting environmental degradation over time, a 'top-down' model. Its solution is a scaled logistic pdf, hence the logistic function is fitted directly to data. The models are illustrated by fitting them to data on the soybean and greenbug aphids. The model derivations and fittings are interesting exercises on 'real-world' applications of mechanistic models and on novel uses of pdfs.

An Improved Method for Retrospective Analysis of Multivariate Individual Observations

✱Joe H. Sullivan, Mississippi State University, MS 9582, MSU, MS 39762, *ymail397-gen9@yahoo.com*

Key Words: robust estimation, statistical process control (SPC)

The analysis of a fixed number of multivariate individual observations for step shifts in the statistical distribution of the observations, possibly combined with multiple outliers, depends on an accurate estimate of the covariance matrix. This research offers a method for partitioning the observations so that within each part the observations are homogeneous while adjacent parts have statistically significantly different mean vectors.

Monitoring Simultaneously the Mean Vector and Covariance Matrix in Process Industries

✱John C. Young, McNeese State University, 1750 Bilbo St., Lake Charles, LA 70601, *johncyoung@bellsouth.net*; Robert L. Mason, Southwest Research Institute; Youn-Min Chou, University of Texas at San Antonio

Key Words: Generalized Variance, Scatter Matrices, T-Square Statistic, Wilks Ratio Statistic

In this paper, we propose to simultaneously monitor the mean vector and covariance matrix of a multivariate normal process using two separate control statistics. One statistic checks for changes in the mean vector and the other for changes in the covariance matrix. The proposed procedure will readily detect the appropriate signal in three cases: (1) the mean vector shifts without a shift in the covariance matrix, (2) the covariance matrix shifts without a mean vector shift, and (3) both the mean vector and covariance matrix simultaneously shift as the result of a change in some key process variables. An advantage of this procedure is that it does not require that the number of new observations exceed the number of process variables.

Using Stepwise Discriminant Analysis as a Post-Hoc Procedure to a Significant Hotelling's T2 Control Chart

Jay Schaffer, University of Northern Colorado; ✱Emily Pollock, University of Northern Colorado, ASRM, McKee 530, Greeley, CO 80639, *emypollock@hotmail.com*

Key Words: Multivariate QC, Discriminant Analysis, Hotelling's T2

This study analyzes stepwise discriminant analysis as a follow-up to a significant Hotelling's T2 control chart using squared partial correlations. Different levels of monitored variables, effect sizes, correlations, sample sizes, and alpha levels are considered. Due to its multivariate nature, it is hypothesized that the results of this study may demonstrate that stepwise discriminant analysis using squared partial correlations is an effective and accurate follow-up procedure for statistical process control (SPC). In addition, the results of this study may result in recommendations surrounding number of monitored variables, effect sizes, correlations, sample sizes, and alpha levels for practitioners using stepwise discriminant analysis as a post hoc procedure to a significant Hotelling's T2 in practice.

Development of Chart Parameters and ARLS for the Expansion of the MEWMS and MEWMV with Individual Observations

Jay Schaffer, University of Northern Colorado; ✱Chad Eshelman, University of Northern Colorado, ASRM, McKee 530, Greeley, CO 80639, *chad.eshelman@gmail.com*

Key Words: Multivariate QC, MEWMS, MEWMV

In SPC, continued development of techniques look for new monitoring charts for processes with multiple correlated variables. Two such charts are the multivariate exponentially weighted moving standard deviation (MEWMS) and multivariate exponentially weighted moving variance (MEWMV). Originally developed by Huwang, Yeh, and Wu (2007), these control charts monitor the trace of the respective covariance matrices for a change in values of the multivariate process using individual observations. Originally, control chart parameters were developed during the simulation process for p = 2 and p = 3 process variables. Using computer simulations of 20,000 replications, further development of the MEWMS and MEWMV used p = 5 and p = 10 correlated variables with individual observations to develop control chart parameters and derive average run lengths similar to those derived by Huwang, Yeh, and Wu.

Reduction in Residential Energy Consumption

✱Seema Seema, McNeese State University, Department of Engineering, P.O. 91735, Lake Charles, LA 70609-1735, *seema.seema85@yahoo.com*; Rebekah A. Griffith, McNeese State University; D. John Griffith, Jr., McNeese State University

Key Words: Hotelling's T-Square, Energy Conservation, Electric Power Usage

Energy data for home and apartment use is collected over heating and cooling cycles to develop an historical data set. Variables measured include day and night electrical power consumption, household activity, indoor set-point temperatures, minimum and maximum daily environmental temperatures, and daily precipitation. A multivariate model is constructed using both regression and Hotelling's T2 analysis. Improvements are made to reduce energy consumption due to heat gain / loss to and from the residence. Further data is then collected during both heating and cooling periods. The effects of the improvements are detected with the T2 analysis and quantified by the estimated parameters of the altered regression model. Payback on the energy improvements is quantified from the statistical model, and compared to the payback expected from engineering considerations alone.

Start-Up Demonstration Tests Using Scans and Intermediate Performance Levels

✱William S. Griffith, University of Kentucky, Department of Statistics, University of Kentucky, Lexington, KY 40506-0027, *griffith@ms.uky.edu*; Michelle Smith, Eastern Kentucky University; Ken Dutch, Eastern Kentucky University

Key Words: Quality Control, Start-up Demonstration Test, Scan, Multistate

Start-up demonstration tests are used to decide whether to accept or reject equipment. A number of tests have been studied in the literature. In this paper, we look at some test criteria which involve the number of successes or failures within a specified number of trials, rather than consecutive successes or failures. We also look at ways that incorporate the notion of intermediate levels of performance or types of successes and failures. Some use will be made of Markov chain methodology which has been used by the authors previously to study other types of start-up demonstration tests.

Sensitivity to Distance, Baseline Distributions, and Forecast Evaluation in Decision and Risk Analysis

✱Victor R. Jose, Duke University, Box 90120, Duke University, Durham, NC 27708-0120, *vrj@duke.edu*; Robert Nau, Duke University; Robert Winkler, Duke University

Key Words: Probability elicitation and verification, Scoring rules, Probabilistic risk analysis, Expert opinion

Assessing the quality of probabilistic forecasts through scoring rules as been an important area of study in the field of decision and risk analysis. In this paper, we construct rich families of scoring rules that are strictly proper (thereby encouraging truthful reporting), sensitive to distance (thereby taking into account ordering of events), and incorporate a baseline distribution relative to which the value of a forecast is measured. We utilize the intuition from the creation of the ranked probability score and extend it from being based on a quadratic rule to any strictly proper scoring rule. In particular through the use of the power and pseudospherical scoring rules, we can create a rich family of scoring rules with the said properties. Areas of possible applications include engineering, meteorology, economics and psychology.

Contracts and Performance in Major League Baseball

✱Elaine Allen, Babson College, Arthur M Blank Center for Entrepreneurship, Blank 102, Babson Park, MA 02457, *allenie@babson.edu*; Julia Seaman, Pomona College

Key Words: baseball, data mining, contract year, batting performance

A baseball player's day-to-day performance is largely unpredictable. For example, a player could have an excellent game one day and go hitless the next. However, by looking at statistics on an annual basis, trends can be identified and investigated. The purpose of this paper is to investigate a purported trend in Major League Baseball that players in their "walk year," meaning that the player is not signed to a contract beyond the end of the current season, will have a "career year" of better than average statistics. The implication is that the player is giving more effort and performing better with their next contract on the line. In order to frame the investigation, only batting performance statistics will be used in this statistical and data mining analysis. Data from 1960 through 2007 will be utilized in this analysis.

Simulating a College Basketball Universe

✱Robert B. Davis, Miami University, 1601 University Boulevard, Hamilton, OH 45040, *davisrb@muohio.edu*

Key Words: simulation, games, mathematical modeling

Many statisticians and programmers have written simulation programs to simulate sports events based on the statistics of the participants. However, this problem becomes much more complicated when one is simulating not simply one game, but attempting to encompass the entirety of a sport. This poster will present an overview of the many challenges that must be faced and the questions that must be answered in order to create a complete and convincing simulated college basketball universe, in which everything from recruiting and in-game tactics to a national tournament and an NBA draft are included. The issues faced in simulating games based on traits of fictitious players rather than actual statistics of real players will also be addressed.

SiSSYS: Student Work from a Capstone Course Using Sports Data - Part I

✱Robin Lock, St. Lawrence University, Dept of Math, CS and Stats, Canton, NY 13617, *rlock@stlawu.edu*; Travis Atkinson, St. Lawrence University; Michael Schuckers, St. Lawrence University

Key Words: capstone, sports, education

During the 2007–08 academic year, a capstone course on the statistical analysis of sports data was conducted at St. Lawrence University. Eight students were involved in this seminar, known as the Statistics in Sports Senior Year Seminar (SiSSYS). In this presentation, we summarize some of the analyses produced by these students. These include applications of previously described statistical work in the literature to current data sets in different sports. As part of this poster we discuss the data and methods utilized by students in this seminar. Another talk at this JSM will discuss the organization and implementation of such a capstone experience.

SiSSYS: Student Work from a Capstone Course Using Sports Data - Part II

✱Travis Atkinson, St. Lawrence University, Dept of Math, CS and Stats, Canton, NY 13617, *tatkinson@stlawu.edu*; Robin Lock, St. Lawrence University; Michael Schuckers, St. Lawrence University

Key Words: sports, capstone, education

During the 2007–08 academic year, a capstone course on the statistical analysis of sports data was conducted at St. Lawrence University. Eight students were involved in this seminar, known as the Statistics in Sports Senior Year Seminar (SiSSYS). In this presentation, we summarize some of the analyses produced by these students. These include applications of previously described statistical work in the literature to current data sets in different sports. As part of this poster we discuss the data and methods utilized by students in this seminar. Another talk at this JSM will discuss the organization and implementation of such a capstone experience.

Determination and Analysis of Factors Determining the Outcomes of National Football League Games

✱Christopher Cohea, Oklahoma State University, Department of Statistics, Stillwater, OK , *statboy78@gmail.com*; Mark E. Payton, Oklahoma State University

Key Words: Football, Sports, NFL, Logistic, Regression, Factor Analysis

The outcome of a National Football League game depends on numerous factors related to events that occur during the game. Over 40 variables were collected on a sample of 534 NFL games from the 2004 and 2005 seasons. Data from only one team (selected at random) will be utilized for each game. Logistic regression is employed to assist in understanding the relative importance of the collected variables. Due to multicollinearity among the predictor variables, factor analysis was used to determine if a set of underlying characteristics exists that explains a significant portion of the measured variability. The factors are named according to their perceived characteristics and are used as predictor variables in a logistic regression model. Game outcome predictions are also explored with both the logistic model with factors and the logistic model with the original variables.

Analysis of Motorcycle Grand Prix Lap Times

✱Leanne Streja, University of California, Los Angeles, School of Public Health, Department of Biostatistics, Los Angeles, CA 90095-1772, *lstreja@ucla.edu*; Robert E. Weiss, University of California, Los Angeles; Catherine A. Sugar, University of California, Los Angeles

Key Words: sports, graphics, longitudinal, repeated measures

Motorcycle Grand Prix Racing (MOTOGP) is the most elite class of motorcycle racing, with 16 to 18 races held annually at race tracks worldwide. Tracks are unique in length and layout with various tight turns and straight-aways. Lap time provides information on a rider's ability, consistency and bike performance. We graphically analyze MOTOGP lap time data from 2002-2007 to assess differences across riders, tracks, and years. Interesting findings include: clustering of rider positions, with considerable separation in race time among the top 5-10 from the rest, lower variability in lap times for the top ten, and determination of a race winner midway through the race, where the gap in time increases substantially between the leader and the rest of the racers. These observations across riders, tracks, and years can be used to adjust a rider's strategy to win the race.

454 COPSS Awards and Fisher Lecture

ASA, ENAR, IMS, SSC, WNAR, Committee of Presidents of Statistics Societies, Biometrics Section

Wednesday, August 6, 4:00 p.m.–5:50 p.m.

Presentation of Awards

✱Jessica Utts, University of California, Irvine, Irvine, CA 92697, *jmutts@ucdavis.edu*

The President's Award recognizes outstanding contributions to the profession of statistics by a young member of one of the participating societies. Sponsored and presented annually by COPSS. The Elizabeth L. Scott Award recognizes a statistician who has fostered opportunities in statistics for women. It honors the lifelong efforts of its namesake, a statistician and astronomer. The award recognizes excellence in: developing programs to encourage women to seek careers in statistics; successfully mentoring women students or new researchers; working to identify gender-based inequities in employment; and serving as a role model to women. It was established in 1992 and is awarded biennially. Fisher Lecturer was established by COPSS in 1963 to honor the contributions of Sir Ronald Fisher and the work of a present-day statistician's contributions to the collection and interpretation of data.

The Population Science Research Agenda and the Women's Health Initiative

✱Ross L. Prentice, Fred Hutchinson Cancer Research Center/The University of Washington, 1100 Fairview Ave. North, M3-A410, Seattle, WA 98109, *rprentic@whi.org*

Key Words: Population Science, Multivariate failure times, Epidemiology, Disease prevention, Nonparametric

Sir Ronald Aylmer Fisher was a central founder of modern statistical science, with multiple pivotal contributions to study design, as well as to data analysis and data interpretation methodology. Crucial current issues for chronic disease population science research include the study designs needed to obtain reliable information on preventive intervention effects, and the adequacy of traditional sources and methods for preventive intervention hypothesis development. Settings in which both randomized clinical trial data and cohort study data are available for the same 'intervention' provide an opportunity to examine the study design issue. The newer types of high-dimensional molecular data that are coming available, especially plasma proteomic data, provide an opportunity to invigorate the preventive intervention development enterprise. Further statistical methods developments are needed.

455 Time Series Analysis via Mechanistic Models ●

Biometrics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Fitting Embedded Population Dynamics Models to Animal Abundance Time Series Data

✱Ken B. Newman, U.S. Fish and Wildlife Service, 4001 N. Wilson Way, Stockton, CA 95205, *ken_newman@fws.gov*; Carmen Fernandez, Instituto Espanol de Oceanografia; Len Thomas, University of St Andrews; Steve Buckland, University of St Andrews

Key Words: state space models, MCMC, particle filters, sequential importance sampling

Monitoring fish and wildlife population abundances and estimating parameters which influence abundance, such as survival probabilities or fecundity rates, are ubiquitous objectives for ecological scientists and resource managers. Bayesian hierarchical models, particularly state space models, are a framework for combining mechanistic population dynamics models with animal abundance time series data as well as accounting for environmental, demographic, and sampling variation. Markov chain Monte Carlo and particle filters are two approaches to fitting such models. Results of a comparative study of the two fitting procedures as applied to grey seal and Chinook salmon data are presented.

Discovering Hidden Latent Components in Continuous-Time Dynamical Systems

✱Giles Hooker, Cornell University, Dept. Bio. Stat. and Comp. Bio, 1186 Comstock Hall, Ithaca, NY 14853, *gjh27@cornell.edu*

Key Words: Dynamics, Neural dynamics, latent components, attractor reconstruction, chaotic data analysis

When investigating the fit of a dynamic mechanistic model, besides misspecification of the terms in the model, it is also possible to consider the addition of extra dynamical components to it. This talk demonstrates the use of techniques from attractor reconstruction to develop tests for model misspecification in the form of missing dynamical components. We proceed by estimating lack of fit in the form of latent influences on the system and then test whether these exhibit dynamical behavior beyond the current system. We outline methods to formalize inference for these diagnostics and to distinguish parameterized models for stochastic noise from latent components. We illustrate these methods with examples from neuroscience where latent components may be interpreted as ion channels across a neuron membrane.

A Bayesian SEIR Approach to Modeling Epidemics

✱Vanja Dukic, The University of Chicago, 5841 S Maryland Ave, MC2007, Chicago, IL 60637, *vanja@uchicago.edu*; Greg Dwyer, The University of Chicago; Bret Elderd, The University of Chicago

Key Words: Epidemics, SEIR, Spatial, Stochastic, vaccination, smallpox

Recent U.S. public policy debates regarding smallpox vaccination were largely focused on comparing mass versus trace vaccination strategies; namely, whether to vaccinate the entire population or only those who have been in contact with infected individuals. In this talk, we present a Bayesian susceptible-exposed-infected-recovered (SEIR) model and apply it to analyze a set of eight smallpox epidemics in Southwest Native American communities during 1780–1781. The outcome of the model is the posterior distribution of epidemic parameters, after taking into account the population and geographical heterogeneity. We then present a comparison of the two main vaccination strategies based on the posterior predictive distribution of the fatalities under each.

456 Reducing Disclosure Risks While Reaching Out to the Data Needs of the Public ●▲

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Releasing Microdata: Disclosure Risk Estimation, Data Masking, and Assessing Utility

✱Natalie Shlomo, University of Southampton, Highfield, Southampton, SO17 1BJ United Kingdom, *N.Shlomo@soton.ac.uk*

Key Words: Disclosure risk measures, Information Loss measures, Disclosure risk - Data Utility trade-off

In order to make informed decisions when releasing microdata, data suppliers must assess disclosure risk. We assume that the microdata contain individuals investigated in a survey and that the population is unknown. The disclosure risk is a function of both the population and the sample, and in particular the cell counts for a contingency table spanned by identifying discreet key variables (i.e., sex, age, occupation, etc.). The disclosure risk measures are estimated using probabilistic modeling. Based on the disclosure risk assessment, data suppliers must choose appropriate SDC methods depending on user requirements, the contents of the data, and the impact on information loss. We develop information loss measures to quantify the effects of SDC methods on statistical analysis. We demonstrate the risk-utility assessment on samples drawn from a census and real data sets.

Web-Based Data Query Tools: Meeting User Needs

✱J. Neil Russell, National Center for Education Statistics, 1990 K. St., NW, Rm. 9050, Washington, DC 20006, *neil.russell@ed.gov*

Key Words: Confidentiality, Disclosure, Web Based Tools, Data User Needs

Some National Center for Education Statistics (NCES) data are only released as restricted-use files and/or through NCES' web-based Data Analysis System (DAS). The DAS allows the user to create programming files, tables of estimates, regression models, and view output. Because of this, there has been a recent policy shift away from PUMFs and towards the DAS as the main means for releasing public-use data. This policy shift has implications for the DAS: 1) meeting data user needs for increased access to data, 2) maintaining the confidentiality of the data, and 3) increasing the number and type of statistical methods provided for the user. A NCES task force was recently charged with examining the DAS data protection methods versus users' statistical and data needs. Some of the task force findings will be presented with a description of how the DAS currently protects data.

New Ways To Provide More and Better Data to the Public While Still Protecting Confidentiality

✱Laura V. Zayatz, U.S. Census Bureau, Commerce/Census/SRD, 4600 Silver Hill Road, Washington, DC 20233-9100, *laura.zayatz@census.gov*

Key Words: disclosure avoidance, data dissemination

The U.S. Census Bureau collects its survey and census data under Title 13 of the U.S. Code which promises confidentiality to its respondents. The agency also has the responsibility of releasing data for the purpose of statistical analysis. In common with most national statistical institutes, our goal is to release as much high quality data as possible without violating the pledge of confidentiality. We apply disclosure avoidance techniques prior to publicly releasing our data products to protect the confidentiality of our respondents and their data. This paper discusses three areas of current disclosure avoidance research: noise for tabular magnitude data, synthetic tabular frequency and microdata, and a remote microdata analysis system.

457 Analysis and Sample Size for Cluster Samples: Community-Based and Group-Randomized Designs ●▲

ENAR, Section on Survey Research Methods, WNAR, Biometrics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Estimation and Inference for Group-Randomized Trials with a Binary Outcome

✱Sherri L. Pals, Centers for Disease Control and Prevention, MS E45, 1600 Clifton Road, Atlanta, GA 30333, *sfv3@cdc.gov*

Key Words: Group-randomized trials, Intraclass correlation, Binary variables

Group-randomized trials (GRTs) require design and analytic methods that take into account intraclass correlation due to assignment of intact social groups to study conditions. Binary outcome variables in GRTs are common, and investigators have increasingly responded to the challenge of developing methods for these trials. Early GRTs with binary outcomes were often analyzed with two-stage methods, involving analysis of group-level event rates. Generalized estimating equations (GEE) provided a method for taking intraclass correlation into account while avoiding the specification of the joint distribution; recently proposed corrections have addressed GEE's poor performance in small GRTs. The use of mixed-model approaches has increased dramatically due to availability in commercial software. Each of these methods will be reviewed, highlighting recent developments and areas for future work.

Power for Clustered Gaussian Data

✱Jacqueline L. Johnson, Novartis Pharmaceuticals, One Health Plaza, Building 405, Room 4031B, East Hanover, NJ 07936, *jackie.johnson@novartis.com*; Keith E. Muller, University of Florida; Diane J. Catellier, The University of North Carolina at Chapel Hill; Mildred M. Maldonado-Molina, University of Florida; Kelli A. Komro, University of Florida

Key Words: group randomized trials

Even though the data analysis in most cluster studies with Gaussian data uses some type of weighting to account for imbalance, current methods for computing power assume equal cluster sizes, for which exact power can be computed. Unbalanced data typical of cluster designs can make actual power differ substantially from the planned power. We describe how to align power calculations with a hybrid method of data analysis which combines mixed model estimation of intraclass correlation with exact transformation to a univariate model for approximate hypothesis testing. Analytic results, enumerations, and simulations all support the approach. The method allows accounting for within-cluster covariates. A power analysis for a study of adolescent drinking behavior illustrates how easily the method can be implemented with standard data analysis and power software.

Estimation and Inference for Clustered Gaussian Data

✱Diane J. Catellier, The University of North Carolina at Chapel Hill, 137 E Franklin Ste 203, CB #8030, Chapel Hill, NC 27599, *diane_catellier@unc.edu*; Jacqueline L. Johnson, Novartis Pharmaceuticals; Keith E. Muller, University of Florida

Key Words: group randomized trials

Gaussian clustered data are often analyzed using a mixed model on individual data or least squares regression of cluster means. Both techniques provide unbiased hypothesis tests for data when the number of clusters and number of observations within each cluster is large. In small samples with unbalanced data, even moderate imbalance in cluster size across treatment groups can lead to test size bias. The magnitude of bias in type I error is compared for the mixed model analysis and weighted and unweighted least squares analysis of cluster means. Of the 10 methods considered, the analysis of cluster means with means weighted by their estimated theoretical variance best controlled type I error. Several methods provided unbiased inference when treatment groups were close to balanced in the number of clusters and observations per cluster, as is often common in group randomized trials.

Internal Pilot Designs for Cluster Samples

✱Christopher S. Coffey, The University of Alabama at Birmingham, 309C RPHB, 1665 University Boulevard, Birmingham, AL 35294-0022, *ccoffey@uab.edu*; Matthew J. Gurka, University of Virginia; Keith E. Muller, University of Florida

Key Words: Sample Size Re-Estimation, Adaptive Designs, Power

Internal pilot designs allow revising the variance estimate at an interim stage to adjust the final sample size up or down, as needed, to achieve the power desired. Clinical trialists have begun using such designs with univariate outcomes. However repeated measures make an internal pilot difficult because the nuisance parameter becomes a covariance matrix rather than a single variance. Compound symmetry within clusters allows applying Gurka, Coffey, and Muller's (2007) exact method for internal pilots with equal cluster sizes. For unequal cluster sizes, an easy to implement hybrid approach uses mixed model estimation of intraclass correlation and exact transformation to an inference-equivalent univariate model. Estimated weights allow using existing univariate internal pilot techniques. Simulations and enumerations demonstrate the accuracy and benefits of the approach.

458 Data Depth–Based and Related Nonparametric Multivariate Procedures ●▲

Section on Nonparametric Statistics, IMS
Thursday, August 7, 8:30 a.m.–10:20 p.m.

Nonparametric Outlier Identification in Multivariate and Functional Data Settings and Beyond

✱Robert Serfling, University of Texas at Dallas, Richardson, TX 75080, *serfling@utdallas.edu*

Key Words: Nonparametric, Projection pursuit, Multivariate, Outlier identification

Broadened foundations and criteria for outlyingness functions – for location, dispersion, regression, and arbitrary "parameter" and data, are discussed. Unifying structures for univariate location outlyingness, "substitution" and "projection pursuit" approaches for multivariate extensions, and masking and swamping breakdown points, are described. For the general case, a "fitting-residuals paradigm" identifies outliers in data via residuals analysis after fitting a suitable parameter. Here a data-based outlyingness function is constructed on parameter space as a function of evaluations at data points of a real-valued nonnegative "criterion function" defined on data space and parameter space and increasing with unfitness of the parameter value relative to the data point. General approaches to such constructions are formulated and illustrative applications described.

Data Depth–Based Inference Procedures That Outperform Classical T (Or T^2 in High D) Procedures

✱Yijun Zuo, Michigan State University, A440 Wells Hall, East Lansing, MI 48823, *yijun.zuo@gmail.com*

Key Words: Data depth,, inference procedures, classical t-procedure, Hotelling's T^2 procedure, boot-strap

In this talk, multidimensional trimming based on "data depth" is discussed. It is found that multi-dimensional depth-trimmed means can possess very desirable properties such as high efficiency and high robustness. Further, inference procedures based on the depth-trimmed means can outperform the classical t-procedure in one dimension and Hotelling's T^2 one in high dimensions.

Data Depth–Based Nonparametric Methods for Multivariate Right-Censored Observations

✱Shojaeddin Chenouri, University of Waterloo, Department of Statistics and Actuarial Science, University of Waterloo, 200 University Ave. W., Waterloo, ON N2L 3G1 Canada, *schenour@uwaterloo.ca*

Key Words: Data depth, multivariate survival functions, multivariate right-censored observations, nonparametric methods, estimation, hypothesis testing

In this talk, we are concerned with nonparametric inferences about multivariate survival function of p dependent survival times under random right censoring. The procedures we develop are based on the notion of data depth. We introduce asymptotically distribution free test statistics for comparing two or more multivariate survival functions. In addition, we propose a nonparametric method to estimate the multivariate survival function.

459 Statistical Aspects of Regional Climate Modeling ▲

Section on Statistics and the Environment, WNAR
Thursday, August 7, 8:30 a.m.–10:20 a.m.

The North American Regional Climate Assessment Program (NARCCAP): Overview and Early Results

✱Linda O. Mearns, National Center for Atmospheric Research, P. O. Box 3000, Boulder, CO 80307, *lindam@ucar.edu*

Key Words: Design of Modelling Experiments, Global Climate Modals, Regional Climate Models, Uncertainty of Climate Model Results

NARCCAP is an international program that is serving the climate scenario needs of the United States, Canada, and northern Mexico. We are systematically investigating the uncertainties in regional scale projections of future climate and producing high resolution climate change scenarios using multiple regional climate models (RCMs) and multiple global model responses to a future emission scenario, by nesting the RCMs within atmosphere ocean general circulation models (AOGCMs) forced with the A2 SRES scenario, over a domain covering the conterminous US, northern Mexico, and most of Canada. The project also includes a validation component through nesting the participating RCMs within NCEP reanalyses.In this overview talk, results from Phase I of the project, the RCM simulations using boundary conditions from NCEP reanalyzes, will be presented. In addition, outcomes of a combined modeler and user group meeting for NARCCAP will be discussed.

Models and Models: Combining Regional Climate Model Output via Multivariate Spatial Models

✱Stephan Sain, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000, *ssain@ucar.edu*

Key Words: Climate change, Markov random field, Bayesian hierarchical models

The North American Regional Climate Change Assessment Program (NARCCAP) is an ambitious experiment that seeks to use a number of regional climate models (RCMs) to produce a wide range of model output over North America. Our goal within this program is to develop statistical methodology

to analyze this model output and assess the sources of uncertainty. To that end, we are developing a Bayesian hierarchical framework that is based upon a multivariate spatial model. This allows us to capture the complex distribution of the spatial fields produced by these regional climate models and make inferences about the model differences as well as projections of climate change. In this talk, the methodology will be discussed and examples of the implementation presented. Examples based on initial NARCCAP runs and other sources will be examined.

Investigating Concurrently High Values of Convective Available Potential Energy (CAPE) and Wind Shear in Global Reanalysis and Climate Model Output

✱Eric Gilleland, National Center for Atmospheric Research, Research Applications Laboratory, 3450 Mitchell Lane, Boulder, 80301, *ericg@ucar.edu*

Much interest concerning climate change has recently been focused on extremes of various atmospheric processes (e.g., severe storms, tornadoes, hail, heavy precipitation, etc.). Many of these processes occur at very fine spatial scales; too fine for current climate models with much coarser spatial support. One approach to gleaning information about such extremes under a changing climate is to study large-scale indicators of such events. Concurrently high values of convective available potential energy (cape) and 0-6 km wind shear have been found to be associated with more intense storms. The present study investigates these variables using current global reanalysis data as well as some climate model output to investigate how this indicator is changing over time by way of statistical models.

460 Statistical Measures Can Help Restore Confidence in U.S. Elections ●▲

Social Statistics Section, Section on Government Statistics, Section on Survey Research Methods, Scientific and Public Affairs Advisory Committee
Thursday, August 7, 8:30 a.m.–10:20 a.m.

National Election Scorecard

Fritz Scheuren, The University of Chicago; ✱Edward Mulrow, National Opinion Research Center, 4350 East-West Highway, Suite 800, Bethesda, MD 20814, *mulrow-edward@norc.org*

Typically, in every recent election, we have been inundated with polling results. The focus of these has been on the potential winning candidate and who will or did vote for him or her. Seldom have there been surveys that asked voters about the voting experience, itself. This paper discusses the first national attempt at such a poll. Issues of design are covered, with sampling and especially nonsampling issues given prominence. There have been smaller efforts of this sort in 2004 and 2006, but not a national attempt. Our focus then will be on what in other contexts might be called a national customer satisfaction survey. Key here, of course, is obtaining the interviews before the winners are announced.

Voter Confidence and the Election Day Voting Experience

✱J. Quin Monson, Brigham Young University, Dept. of Political Science, 745 SWKT, Provo, UT 84602, *Quin.Monson@byu.edu*; Kelly Patterson, Brigham Young University; David Magleby, Brigham Young University; Ryan Claassen, Kent State University

Key Words: voter satisfaction, election administration, voter confidence, exit polls, elections, voting

The underlying objective of voting reforms is to improve public confidence in the election system; however, little is known about the conditions voters actually face and how this affects their confidence. In 2006, the Center for the Study of Elections and Democracy at Brigham Young University collected exit poll data on the voting experience and characteristics of individual voters together with data of the actual conditions at the polling locations in Franklin (Columbus) and Summit (Akron) counties in Ohio as well as statewide in Utah. The findings reveal good reasons for boards of elections to invest in improving service at polling places. The service voters receive at the polling place affects their confidence that ballots will be counted accurately. Equipment differences across jurisdictions, wait times, voter sense of privacy, and voter partisanship also affect confidence.

How ASA Members Are Helping States Improve Elections

✱Arlene S. Ash, Boston University, 801 Massachusetts Avenue, Second Floor, Boston, MA 02118, *aash@bu.edu*

Key Words: sampling, voting, audits

Many states are currently writing or revising laws for election auditing. ASA members have been actively working with voting rights activists, computer scientists, state legislators, county supervisors of elections and secretaries of state to ensure the use of credible audit procedures for elections. Issues range from highly pragmatic (e.g., how to randomly select precincts in a way that is credible to public observers, how to efficiently and accurately tally votes from multiple races) to more theoretical (e.g., initial sampling designs, efficient sequential procedures for when findings from the initial sample are equivocal, post-hoc power calculations). The work is important, and the window of opportunity to improve all aspects of electoral process is open.

461 Symbolic Data: Theory and Methods ●▲

IMS, Section on Nonparametric Statistics
Thursday, August 7, 8:30 a.m.–10:20 a.m.

Symbolic Data Examples, Analytic Aspects, and SODAS Software

✱Edwin Diday, Paris Dauphine University, PL; Mle De Lattre de Tassigny, Paris, 75016 France, *diday@ceremade.dauphine.fr*

Key Words: Data Mining,, conceptual statistics, Symbolic Data Analysis,, Knowledge mining,

Databases are now ubiquitous in industrial companies and public administrations and they often grow to an enormous size. They contain units described by variables that are often categorical or numerical (which can be also transformed to categories). It is then easy to build categories or Cartesian product of categories or categories by using a clustering process which yields to clusters defining each category. Symbolic data represented by structured variables, intervals, list, histograms, distributions, curves and the like, keep the "internal variation" of categories better than do standard data. The aim of Symbolic Data Analysis is to generalize Data Mining and Statistics to higher-level units called "concepts" (which represents these categories) described by symbolic data. The SODAS software (sponsored by EUROSTAT) extends standard tools of Statistics and Data Mining to these units.

Mixture Modeling of Symbolic Data

Matthieu Vrac, Ecole Polytechnique; Edwin Diday, Paris Dauphine University; ✱Lynn Billard, The University of Georgia, Department of Statistics, Athens, GA 30602-1952, *lynne@stat.uga.edu*

We consider the situation where the observations consist of a vector of empirical distribution functions estimated from a larger set of point observations. We develop a methodology for separating these distributions into clusters based on a probalistic law, using a mixture of copula functions. The results are applied to a set of meteorological data.

Regression Models for Symbolic Interval-Valued Data

✱Francisco de A.T. De Carvalho, Centro de Informatica - CIn/UFPE, Av. Prof. Luiz Freire, s/n - Cidade Universitaria, Recife-PE, 50740540 Brazil, *fatc@cin.ufpe.br*; Eufrasio de A. Lima Neto, Centro de Informatica - CIn/UFPE

Key Words: Symbolic Data Analysis, Symbolic Interval-Valued Data, Regression

This presentation concerns regression models for symbolic interval-valued data. In the framework of Symbolic Data Analysis (Billard and Diday (2007)), Billard and Diday (2000) presented an approach to fitting a linear regression model on interval-valued data based on the information given by the midpoints of the intervals. Lima Neto and De Carvalho (2008) improved the former approach presenting a method based on the information given by the midpoints and ranges of the intervals. However, none of these regression models are able to guarantee that the predicted values of the lower bounds will be lower or equal than the predicted values of the upper bounds of the independent variable. In order to guarantee this kind of constrains, we considerer here constrained linear regression models to interval-valued data based on the information given by the midpoints and ranges of the intervals.

462 *Statistics in Biopharmaceutical Research*: The ASA's New Journal ●▲

Statistics in Biopharmaceutical Research Journal, Biopharmaceutical Section, WNAR

Thursday, August 7, 8:30 a.m.–10:20 a.m.

A Conditional T Suite of Tests for Identifying Differentially Expressed Genes in a DNA Microarray Experiment with Little Replication

✱Dhammika Amaratunga, Johnson & Johnson Pharmaceutical R&D, LLC, *damaratu@prdus.jnj.com*; Javier Cabrera, Rutgers, The State University of New Jersey

Key Words: microarray, t test, borrowing strength, resampling

We will present a method of borrowing strength in a distribution-free framework for analyzing data generated by experiments in which a large number of entities are measured in a few samples. We focus on the comparison of two groups of DNA microarray data to determine which genes are expressed differentially between the two (e.g., in diseased vs. normal tissue). In such experiments the number of genes is in the tens of thousands while the number of samples per group is often fewer than ten. While these data could be simply analyzed gene by gene with a series of individual t tests, the thought of increasing the overall power of the procedure by borrowing strength across genes could be entertained. We will describe one approach, the conditional t procedure, and demonstrate its superiority to the simple t test and to other procedures. We will also discuss extensions to other settings.

Methods for Assessing Improvement in Specificity When a Biomarker Is Combined with a Standard Screening Test

✱Pamela A. Shaw, National Institute of Allergies and Infectious Diseases, 6700A Rockledge Drive, RM 5230, Bethesda, MD 20892, *shawpa@niaid.nih.gov*; Margaret S. Pepe, Fred Hutchinson Cancer Research Center; Todd A. Alonzo, University of Southern California Keck School of Medicine; Ruth Etzioni, Fred Hutchinson Cancer Research Center

Key Words: ROC curve, Relative accuracy, Specificity, Diagnostic tests, Study design

Biomarkers that can be used in combination with established screening tests to reduce false positive rates are in considerable demand. We present methods for evaluating the diagnostic performance of combination tests that require positivity on a biomarker test in addition to a standard screening test. These methods rely on relative true- and false-positive rates to measure the loss in sensitivity and gain in specificity associated with the combination relative to the standard test. These methods are extended to evaluate combinations with continuous biomarker tests by introducing a new statistical entity, the relative receiver operating characteristic (rROC) curve. Inference can be made by applying existing ROC methodology. We illustrate the methods with two examples from cancer screening.

Nonproportional Hazards and the Power of Sequential Tests

✱Qi Jiang, Amgen, Inc, *qjiang@amgen.com*; Steven Snapinn, Amgen, Inc

Key Words: Nonproportional hazards, Lag, Noncompliance, Power, Sequential trials, Futility

While statistical methods exist for the calculation of sample size and the analysis of standard endpoint trials with nonproportional hazards, the impact of nonproportional hazards on the properties of sequential clinical trials has not been well studied. In this paper we use simulation to evaluate the impact of nonproportional hazards on the power of standard group sequential methods. We consider boundaries designed to reject H0 only, H1 only, or both H0 and H1. Key findings include 1) The presence of onset lag causes a greater decrease in the power of a sequential method than the power of the fixed-sample method. The impact is particularly great on sequential methods that allow for early rejection of H1. 2) The presence of noncompliance causes a smaller impact on the power of a sequential method than on the power of the fixed-sample method.

463 Communicating Statistics: Speaking Out and Reaching Out ●▲

Section on Statistical Consulting, Section on Quality and Productivity, Section on Statistics in Epidemiology, Section on Teaching Statistics in the Health Sciences, WNAR

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Acquiring Knowledge and Deep Understanding of an Application Field

✱Stephan Ogenstad, Statogen Consulting LLC, 1209 Winkworth Way, Wake Forest, NC 27587, *sogenstad@Statogen-Consulting.com*

Key Words: Field knowledge, Communication, Statistical scientist, Training, Leadership

Unlike other disciplines, statistics does not develop from itself. It needs motivation from new problems arising in all areas of life. The future of statistics lies in communication of statisticians with other researchers in other areas of learning. Statisticians trained at our universities tend to be more technique-oriented than application-oriented. Most statisticians are not educated thoroughly in application-oriented problem solving in their formal training. Nonetheless, they have powerful tools to solve problems. The world is rapidly becoming quantitative and data focused. Many professions depend on numerical measurements to make decisions in the face of uncertainty. Statisticians need quantitative abilities, statistical knowledge, and communication skills to work in challenging fields.

Communicating Statistics to Nonstatisticians Clearly and Persuasively

✱Lothar T. Tremmel, Cephalon, Moores Road, Frzer, PA 19355, *ltremmel@cephalon.com*

Key Words: Communication, Statistical graphs, Visual processing, Persuasive Speech

When planning a study, statisticians must communicate difficult concepts to nonstatistical experts. Too often, we hide behind abstract concepts or focus on "logos" (logical argument), ignoring "pathos" and "ethos" as the other elements of persuasive speech. Real life case studies will be used to demonstrate these principles. While words must be processed sequentially, good graphs harness the eye's parallel processing power. Important features of good graphs will be elucidated by example. Even numbers can be printed in ways that benefit from parallel visual processing, too. We call this property "glanceability" and show how simple changes may turn confusing output into highly glanceable displays. In summary, statisticians are masters of information, but we need to get better at communication. Important tools toward this goal are persuasive speech and effective data presentation.

Attributes of Effective Statistical Leadership in Organizations

✱Thomas W. Dobbins, Merck & Co., Inc., Biostatistics and Research Decision Sciences (BARDS), 351 N. Sumneytown Pike, North Wales, PA 19454-2505, *thomas_dobbins@merck.com*

Key Words: Statistical Leadership, Communicating statistics

Pharmaceutical industry statisticians are facing significant challenges to our profession, including increased demand on our skills due to greater business constraints (time, resources), increased regulatory complexity, and adversarial medical publication and media environments. Thus, there is a pronounced need for effective statistical leaders to ensure that statisticians and statistical organizations continue to thrive and remain relevant. While good technical skill is a necessary attribute of a statistical leader, it is not, and never has been, a sufficient attribute. In this talk, I will discuss key attributes of effective statistical leadership, provide some examples of current gaps, and offer a few recommendations for how statisticians can develop these leadership skills.

464 Meeting Within a Meeting: Promoting Quantitative Literacy by Helping Mathematics and Science Teachers Improve Their Skills in Teaching Statistics ●▲

Advisory Committee on Teacher Enhancement, Section on Statistical Education

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Promoting Quantitative Literacy by Helping Mathematics and Science Teachers Improve Their Skills in Teaching Statistics: The ASA's Meeting Within a Meeting (MWM) Program

✱Katherine T. Halvorsen, Smith College, Clark Science Center, Northampton, MA 01063-0001, *khalvors@email.smith.edu*; ✱Nari Carter, Brigham Young University, 340 MCKB, Provo, UT 84602, *narij@comcast.net*; ✱Aleisha Coleman, Navigator Pointe Academy, 127 E. 400 N., Provo, UT 84606, *aleishaz@hotmail.com*; ✱Jenna Mortensen, Brigham Young University, 230 TMCB, Provo, UT 84602, *jenmarie87@byu.edu*; ✱Rebecca Nichols, The American Statistical Association, 732 North Washington Street, Alexandria, VA 22314, *rebecca@amstat.org*

Key Words: Quantitative Literacy, Teaching Statistics, K-12, MWM, GAISE

It is widely agreed by educators that students in middle school are at the greatest risk of opting out of quantitative subjects and that it is in middle school where the "I don't do math" attitude emerges. Prompted in part by the Federal "No Child Left Behind" legislation, many states have recently strengthened their K-12 quantitative literacy requirements. However, many teachers have been "left behind" since their quantitative skills need to be brought up-to-date. The mission of ASA's MWM program is to help middle school mathematics and science teachers teach statistics in accordance with their state's education standards for quantitative literacy and with ASA's GAISE recommendations. In this session a panel of program participants will present an overview of MWM and a year-long follow-up program. The panel and audience will discuss ways to improve and to expand the program.

465 Challenges and Opportunities for Using R and Other Software in Introductory and Intermediate Probability and Statistics Courses ●

Section on Statistical Education, Section on Nonparametric Statistics, Section on Teaching Statistics in the Health Sciences

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Using R in Probability and Mathematical Statistics Courses

✱Amy Froelich, Iowa State University, 324 Snedecor Hall, Ames, IA 50011, *amyf@iastate.edu*

Key Words: statistical education, teaching statistical concepts, teaching probability distributions

Statistical computing software packages, such as R, have been used primarily in applied statistics courses at both the undergraduate and graduate level. I have found that incorporating R into the calculus-based probability and mathematical statistics courses can facilitate the instruction of many concepts and principles typically covered in these courses. In this talk, I will present an example of how I have used R in these courses to enable students to visualize probability distributions, to explore the properties of these distributions and their use as approximations to data distributions, and to connect these concepts to estimation and inference for the parameters of these probability distributions.

Introductory and Advanced Statistics with a Spreadsheet Interface and Spreadsheet Tools

✱Richard M. Heiberger, Temple University, Department of Statistics, 203k Speakman, 1810 N. 13 St, Philadelphia, PA 191183717, *rmh@temple.edu*; Erich Neuwirth, University of Vienna

Key Words: R, Excel, Spreadsheet, software

R is one of the best statistics programs currently available. Excel is the most prevalent software used for data storage and interpretation. RExcel integrates the Rcmdr menu system onto the Excel menu bar. This combination puts complete access to the full power of R at the mouse-click level. Data can be transferred from Excel to R (and back). The clickable RExcel menu supplements the R command language, a powerful language which can be daunting for an occasional user. Results from the analyses in R can be returned to the spreadsheet. We discuss our experience teaching introductory courses in statistics with the power of R and the ease of Excel. We discuss some of the technical challenges associated with the development, implementation, and applications of this software combination.

‘Software as a Service’: Providing Open Access to Statistical Software Using ASP.NET

✱Neil Polhemus, StatPoint Technologies, Inc., 11133 Victor Drive, Marshall, VA 20115, *neil@statpoint.com*

Key Words: software, students, ASP.NET, Internet, web service, browser

Recent developments in technologies such as ASP.NET are making it possible to provide access to statistical software for students directly over the Internet. It is no longer necessary that students download and install statistical software. Rather, students can access statistical software “as a service” from within their web browser, regardless of the type of computer they own. Data files may be opened transparently, whether they reside on the student’s machine or in the instructor’s directory on a server. Statistical software configured as a web service can then access that data, perform the desired calculations, and return both numerical and graphical results to the browser. This paper demonstrates the implementation of such techniques using STATGRAPHICS Online as an example.

Experiences Using R To Teach Undergraduate Statistics Courses: Enabling Critical Thinking

✱Philip Turk, West Virginia University, Department of Statistics, 423 Hodges Hall, PO Box 6330, Morgantown, WV 26506, *pturk@stat.wvu.edu*

The author has used the statistical software package R extensively to teach undergraduate statistics courses at major state universities, including elementary and intermediate statistics, probability, sampling, ANOVA, regression analysis, time series, and statistical methods. Using R as a key component of these courses has sometimes been a challenging task. In this presentation, the integration of R into these courses will be discussed and experiences gathered from the classroom will be shared. Examples will be provided that are either interesting or demonstrate the ability of R to enable critical thinking and to teach important statistical concepts. I will discuss student feedback and my own personal observations. Impediments to learning and implementation, time management, and web site maintenance are important considerations that must be considered before the course begins.

466 Nonresponse Bias and Callbacks in Household and Establishment Surveys ●▲

Section on Survey Research Methods, Section on Government Statistics, Social Statistics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Models for Callback Procedures and Mode Effects in Household and Establishment Surveys

✱John L. Eltinge, Bureau of Labor Statistics, OSMR, PSB 1950, BLS, 2 Massachusetts Avenue NE, Washington, DC 20212, *Eltinge.John@bls.gov*

Key Words: Cost structures, Nonresponse bias, Measurement error model, Pattern-mixture model, Selection model, Temporal aggregation

This paper presents a review and synthesis of some previous literature on survey callback procedures and related issues arising from changes in contact and collection modes during the callback process. We examine several standard approaches within a general modeling framework defined by fixed-subpopulation analyses, selection models, pattern-mixture models, and measurement error models. Cost structures, bias-variance trade-offs and temporal aggregation receive special attention.

Efficiency and Bias in Differential Contact Strategies

✱Stephanie Eckman, National Opinion Research Center/JPSM, 4350 East West Highway, Bethesda, MD 20814, *eckman-stephanie@norc.uchicago.edu*; Colm O’Muircheartaigh, The University of Chicago

Key Words: non-response bias, multi-phase sample design, response prediction

Two-phase sampling and other differential contact strategies can help control costs, maintain weighted response rates and reduce non-response bias. When the relative cost of later call-backs is sufficiently high, these strategies may increase the efficiency of the data collection process, as measured by sampling precision per unit cost. When the respondents obtained through these strategies more effectively represent the pool of nonrespondents, we can reduce nonresponse bias. Using call history data from recent NORC face-to-face surveys, we explore the effects of field strategies that tailor contact protocols to predicted likelihood of response, and find strategies that reduce the bias and variance of the estimates for a given data collection budget.

Nonresponse Bias Analysis Using Critical Item-Only Respondents for the National Survey of Recent College Graduates

✱Donsig Jang, Mathematica Policy Research, Inc., 600 Maryland Ave., SW, Suite 550, Washington, DC 20024, *don_sig@yahoo.com*; Xiaojing Lin, Mathematica Policy Research, Inc.; Kelly Kang, National Science Foundation

Key Words: weighting adjustment, regular respondent, response rate

Reluctant respondents due to respondent burden and lack of motivation have resulted in increased data collection costs as survey researchers have attempted to maintain traditionally high response rates. This reluctance may be related to the amount of time it takes to complete an interview in large-scale multi-purpose surveys, such as the National Survey of Recent College Graduates (NSRCG). Recognizing that respondent burden or questionnaire length may contribute to lower response rates, the NSRCG offered its nonrespondents near the end of the data collection period an opportunity to complete a much abbreviated interview consisting of a few critical question items. This paper describes an investigation of critical-item-only respondents and the extent to which they are different from regular respondents.

Properties of Callback Procedures Under Moderate Deviations from Specified Models

✱Randall Powers, Bureau of Labor Statistics, 2 Massachusetts Ave. NE, Washington, DC 20212, *powers_r@bls.gov*; John L. Eltinge, Bureau of Labor Statistics

Key Words: Incomplete data, Logistic regression, Loglinear model, Nonresponse adjustment, Simulation study, Weighting adjustment

For many household and establishment surveys, initial contact and interview attempts by themselves do not produce satisfactory response rates. This often leads data-collection organizations to use various callback procedures to collect data from sample units that initially were nonrespondents. Analyses of the resulting data depend implicitly or explicitly on models for the response mechanism, conditional on the specific callback procedure. These models generally are acknowledged to be, at best, approximations to more complex underlying response processes. Consequently, it is important to explore the extent to which the properties of the callback-adjusted estimators may be sensitive to deviations from the assumed models. This paper develops a framework for exploring the impact of moderate deviations, and presents some related simulation results.

467 Bayesian Issues in Medical Devices

Section on Bayesian Statistical Science, Biopharmaceutical Section, Section on Health Policy Statistics

Thursday, August 7, 8:30 a.m.–10:20 a.m.

A Semiparametric Bayesian Approach to Meta-Analysis of Clinical Trials

✱Lijuan Deng, Boston Scientific Corporation, 100 Boston Scientific Way, M1, Marlborough, MA 01752, *dengl@bsci.com*; Pulak Ghosh, Georgia State University

Key Words: Bayesian, Meta-analysis, Dirichlet Process, mixed-effects, MCMC, WinBUGS

When considering adverse event reporting in clinical trials, it is likely that patients have repeated events in different follow-up times. Previously, to account for the correlation of adverse events of the same patient, a two-level random effects model which assumes Normal distribution of the effects across studies and patients is used. However, the assumption of Normal distribution may result in undue influence on outlying studies as they are shrunk towards the overall mean. To accommodate study-level outliers and hence fit a more flexible distribution than the normal, a Dirichlet process prior is incorporated into Bayesian meta-analysis setting. In-built cluster algorithm and stick-breaking representation are used to model the truncated Dirichlet process which forms the basis of the flexible random-effects distributions. Simulated studies are used to justify robustness of the method.

Assessing the Agreement Between Two Measurements: A Bayesian Alternative to Limits of Agreement

✱Cody Hamilton, Edwards Lifesciences, 92614, *Cody_hamilton@edwards.com*

Key Words: Bayes, Statistical agreement, posterior predictive distribution

This work proposes a predictive interval approach to assessing the validity of a new measurement device by examining its agreement with a method already in general use. The Bayesian posterior predictive distribution may be used to estimate the probability that the difference between two future measurements will be within acceptable limits.

Sample Size: When Bigger Is Not Necessarily Better

✱Jeng Mah, American Medical Systems, Inc., 10700 Bren Road W, Minnetonka, MN 55343, *jeng.mah@ammd.com*

Key Words: Bayesian decision, sample size, minimal loss

Power is a function of sample size. A larger study invariably increases the power of the hypothesis test; only an infinitely large study can result in absolutely correct answers. In any study, some power must be sacrificed to make sample size practical. Nominal power choices (e.g., 80% or 90%) are often used without considering the costs of sampling and the losses incurred from wrong decisions. In a previous study, I proposed a decision model explicitly harnessing loss in clinical trial design and observed that minimal loss can be achieved by choosing decision rules based on posterior probability or the Bayes factor. This paper examines a sample size estimation approach which seeks to minimize the overall expected loss, which is a combination of false positive and negative errors and the cost of sampling. Simple cases with conjugated prior distributions are examined through simulations.

Increasing Power for Early Stopping via Bayesian Prediction in a Medical Device Clinical Trial

✱Andrew Mugglin, The University of Minnesota, 2221 University Ave SE, Minneapolis, MN 55414, *amugglin@umn.edu*

Key Words: Bayesian, clinical trial, medical device, interim analysis, predictive probability

Bayesian clinical trial designs that utilize predictive probabilities for trial adaptations have been successfully used for several years in the medical device regulatory setting. When the time point for the primary outcome is long relative to the enrollment period, certain adaptations (e.g., sample size changes) are difficult. However, by incorporating the use of predictive probabilities based on outcomes that are known early, the power for early stopping can be substantially improved, relative to standard frequentist methods. In this talk we describe a clinical trial design that employs Bayesian methods to predict trial success based on early outcomes. The relationship between predictors and final outcome is learned as the trial progresses. The chief advantage of this approach is an increase in power to stop early, and the time savings can be substantial.

468 Approaches to Maintaining Confidentiality and Privacy in Surveys

Section on Government Statistics, Section on Survey Research Methods, Social Statistics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Total Confusion: Disclosure Risk Assessment for Queriable Web-Based Reporting Systems in a Public Health Context

✱Kevin J. Konty, New York City Department of Health and Mental Hygiene, 836 union st, 3, brooklyn, NY 11215, *kkonty@health.nyc.gov*

Key Words: disclosure risk assesment, tabular data, health policy, public health data, vital statistics, administrative records

Protecting confidentiality is an essential duty of all Health Departments. However, existing guidelines have important limitations, including: 1) poor specification of disclosure, 2) failure to consider multiple data releases, 3) arbitrary rules-of-thumb and 4) conflation of privacy risk and reliability. We address these issues in an increasingly common setting: a queriable web-based reporting system for health events and vital statistics. We compare traditional approaches to more rigorous statistical methods. Two microdata methods (K-anonymity and L-diversity) are applicable but must be modified to address multiple tabular releases, for which we use a guidance matrix (Dandekar). The resulting method is easily applied and in our example verifies that the proposed releases contain no disclosure risk (hospitalizations, births), or only require minor modifications to ensure safety (deaths).

Controlling Disclosure Risk in Synthetic Public Use Microdata: The Longitudinal Business Database

✱Arnold Reznek, U.S. Census Bureau, 4600 Silver Hill Road, CES Room 2K128F, Suitland, MD 20233, *arnold.phillip.reznek@census.gov*; John Abowd, Cornell Univeresity; Saki Kinney, Duke University; Jerome Reiter, Duke University; Javier Miranda, U.S. Census Bureau

Key Words: Confidentiality, Disclosure risk, Synthetic microdata

The U.S. Census Bureau's Center for Economic Studies (CES) and academic researchers are developing synthetic microdata public use files from the Longitudinal Business Database (LBD). The LBD longitudinally links variables for business establishments from the Census Bureau's Business Register annual files. Synthetic microdata sets are designed to mimic the distributions of the underlying "real" microdata data and allow researchers to carry out analytically valid studies, without disclosing confidential information about respondents. The "synthetic LBD" files contain synthesized data on establishments' employment and payroll and on their birth and death years, as well as industry and geography. After briefly describing how we are creating the synthetic LBD and evaluating its analytic validity, we describe the confidentiality protection measures we are applying.

New Developments in the Protection of Confidential Data

✱Sam Hawala, U.S. Census Bureau, 4700 Silver Hill Road, Washington D.C., MD 20233, *sam.hawala@census.gov*

Key Words: Disclosure, Synthetic Data

This paper discusses the combined use of model fitting to data and matching using the model predicted values to reproduce the aggregate behavior and the main features of a data set. We approximate the data using semiparametric regression models combining a simple additive structure with the flexibility of the nonparametric approach. Then we use the model predicted values to simulate new data to preserve confidentiality. In the case of highly sensitive data this method provides the necessary protection from disclosure.

Selective Rounding: A Better Alternative to Conventional Rounding of Tabular Data

✱Ramesh A. Dandekar, Energy Information Administration, 1000 Independence Av, washington DC, DC 20585, *ramesh.dandekar@eia.doe.gov*

Key Words: disclosure control, rounding, tabular data

Multiple variations of "conventional" data rounding procedures have been proposed and are practiced to protect sensitive tabular cells. Irrespective of the inherent procedural variations, all these procedures fall under a general category of "indiscriminant" rounding. The "indiscriminant" rounding procedure results in reduced data utility and quality. To improve the overall utility and quality of the published tabular data, we instead propose using a selective rounding procedure for tabular cells. The proposed procedure could be implemented by using multiple different techniques/tools currently available to the statistical agencies.

Truncated Triangular Distribution for Multiplicative Noise and Domain Estimation

✱Jay J. Kim, National Center for Health Statistics, 3311 Toledo Road, Hyattsville, MD 20782, *pzk3@cdc.gov*; Dong M. Jeong, Korea National Statistical Office

Key Words: masking, confidentiality, truncation, triangular distribution

It has been suggested that the truncated triangular distribution be used for masking microdata to protect the confidentiality of the data. The random variable which follows the truncated triangular distribution serves as a multiplicative noise factor. The plausible candidate distribution for this masking is one which is centered and truncated symmetrically around 1. Note that it is not desirable to use values that are too close to 1 due to confidentiality concerns. The desired functional form of the truncated triangular distribution has been developed by Kim (2007) and used experimentally to mask the Korean Consumer Expenditure Survey data. Domain estimation formulas are developed, so that users can perform domain estimation. In this paper, we will show domain estimation formulas for the data masked by multiplicative noise and some results from the application of this distribution.

469 Time Series, State Space Models, and Other Reflections on the Work of Bob Shumway ●

Section on Physical and Engineering Sciences
Thursday, August 7, 8:30 a.m.–10:20 a.m.

Estimating the Selection Gradient of a Function-Valued Trait

✱Jay H. Beder, University of Wisconsin-Milwaukee, Dept. of Mathematical Sciences, P. O. Box 413, Milwaukee, WI 53201-0413, *beder@uwm.edu*

Key Words: quantitative trait, fitness, selection gradient, function-valued trait, Gaussian process, reproducing kernel Hilbert space

In classical evolutionary biology, a quantitative trait is a finite-dimensional random vector z, whose fitness is modeled by a relative fitness function w(z). The selection gradient is defined to be the vector of coefficients of the best linear predictor of w based on z. This vector is estimated by regressing observed fitness on observed traits. In 1989 Kirkpatrick and Heckman introduced the study of infinite-dimensional or function-valued traits. We will discuss the modeling of such traits, and propose a method for estimating the selection gradient in this setting. Gaussian processes and reproducing kernel Hilbert spaces will play a central role.

Estimation of Multiple Signals from Seismic Array Data

✱Joshua Kerr, California State University, East Bay, 25800 Carlos Bee Blvd., Department of Statistics, Hayward, CA 94542, *joshua.kerr@csueastbay.edu*

Key Words: Multiple Signals, Seismic, Frequency, Azimuth

Estimates of azimuth and velocities are obtained by maximization of a profile-likelihood once the model is transformed to the Frequency Domain. Consistency is shown even when the number of frequencies gets `large' and non-parametric methods are used to find estimates of variances using R.

State-Space Discrimination and Clustering of Atmospheric Time Series Data Based on Kullback Information Measures

✱Joseph E. Cavanaugh, The University of Iowa, Department of Biostatistics, C22 GH, 200 Hawkins Drive, Iowa City, IA 52242-1009, *joe-cavanaugh@uiowa.edu*; Thomas Bengtsson, Bell Labs, Alcatel-Lucent

Key Words: atmospheric science, classification, geostatistics, pattern recognition, state-space process

During his distinguished career, Bob Shumway has been a leading innovator in the development and investigation of methods for time series discrimination and clustering. This presentation is based on research that has directly evolved from his influential work. We consider the problem of identifying geographically homogeneous regions based on similarities in the temporal dynamics of weather patterns. Two disparity measures are proposed and applied to cluster time series of observed monthly temperatures from locations across the state of Colorado. The two measures are based on state-space models, where the monthly temperature anomaly dynamics and seasonal component variation are structurally represented in the latent process. Application of these measures produces clusters consistent with known atmospheric flow patterns.

You Make the Earth Move Under My Feet

✱David Stoffer, University of Pittsburgh, Pittsburgh, PA 15206, *stoffer@pitt.edu*

Key Words: CTBT, Earthquakes

Throughout the late 1990s, U.S. intelligence agencies detected seismic-type events near the former Russian nuclear test site at Novaya Zemlya. Experts were allegedly unable to determine the nature of these events. Unnamed U.S. defense officials later stated that these events looked like covert nuclear tests. Critics of the Comprehensive Nuclear Test Ban Treaty used this incident as proof that the treaty is not verifiable. While those critics clearly lack intelligence, we explore some techniques of discriminating between explosions and earthquakes.

470 Quantifying and Hedging Operational Risk

Section on Risk Analysis
Thursday, August 7, 8:30 a.m.–10:20 a.m.

Business Process-Based Risk Analysis for Enterprise Risk Management

✱Bonnie Ray, IBM T. J. Watson Research Center, 1101 Kitchawan Rd, Yorktown Heights, NY 10598, *bonnier@us.ibm.com*; Lea Deleris, IBM T. J. Watson Research Center; Christian Toft-Nielsen, IBM Business Performance Services; Mihir Shah, IBM Business Performance Services

Key Words: Bayesian networks, enterprise risk management, business process

A key component of enterprise risk management is the ability to systematically identify risks and understand their impacts across an enterprise. From a business perspective, linking risk identification to business process models provides an organizing structure for understanding and quantifying risk interdependencies and their effects. In this paper, we present a formal model for linking risks to business processes and show how the model can be used to provide quantitative analysis for risk mitigation investment decision making under uncertainty. The underlying model is based on the concept of Bayesian networks, expanded to incorporate elements of business process maps. The model and analysis are illustrated using a tool developed to aid in risk assessment for IBM's businesses in emerging markets.

Data- and Model-Oriented Decision Support Systems for Anti-Fraud Automation

✱Raul Moreno, Nara Niva, c Tulipán s/n, 28933 Móstoles Madrid, International Spain, *raul.moreno@naraniva.com*; David Rios Insua, Universidad Rey Juan Carlos

Key Words: Anti-Fraud Automation, Bayesian Network, Transactional Influence Diagram, Complex Event Processing, Telecommunications, Credit

Fraud detection and prevention are increasingly important issues in business risk analysis areas such as credit or telecommunication fraud detection. We provide integration insights from Data Oriented and Model Oriented DSS areas, and investigate on a new hybrid approach in event procesing for solving complex business decisions problems under transactional environments. Data decision support technologies, commercially known as business intelligence, are mainly addressing decision support methods within enterprise environments, we present how Decision Theory and Bayesian Networks, such as Influence Diagram, can improve and automate enterprise decision making for efficient deployment of anti-fraud business systems.

Service Project Risk Estimation

✱Pu Huang, IBM T. J. Watson Research Center, 1101 Kitchawan Rd., Yorktown Heights, NY 10598, *puhuang@us.ibm.com*

Key Words: Risk assessment, Prediction, Information fusion

IT service industry has enjoyed tremendous growth in the past decade. Large IT service companies nowadays typically engage in a large number of projects at any given point of time. Constantly monitoring and managing these projects has become a challenging job, as the pure amount of information is overwhelming. In this talk, we will present an approach that integrates information from diverse sources to predict future project performance. Driven by practical needs, our approach has the following two features: accommodate different time-scales, and auto adapt to prediction error.

Measuring the Impact of IT Capital Investment in Reducing Operational Risk

✱Deborah Cernauskas, IBM, 4541 Saratoga Ave, Downers Grove, IL 60515, *cernausk@us.ibm.com*; Colin Lawrence, FSA

Key Words: operational risk, IT risk reduction

This paper explores the use of a Cobb Douglas Loss production function to capture the impact of IT capital spending in reducing operational losses. The production function relates current IT capital investment spending to operational losses over time. Previous research has shown IT capital investment results in an increase in firm valuation consistent with normal investment returns in the short run. However, it has also been shown that excess returns due to IT investment capital occur over longer time frames. These returns have been attributed to productivity gains and output growth and have ignored operational loss reduction as an alternative explanation. This paper focuses on establishing a relationship between IT capital investment with a reduction in operational losses.

Combining Measures of Risk Across Dependent Series

✱Katherine B. Ensor, Rice University, *kathy@stat.rice.edu*

Key Words: complex risk

Risk metrics are often applied simply to one or a few series or processes at a time. As we well know, the correlation between series can substantially alter the risk landscape. We examine robust strategies to combining risk metrics across multiple dependent series that are not necessarily observed simultaneously. Furthermore, covariates may have a differential impact on the up or downside risk. This modeling complexity is incorporated in our methodologies.

471 Component-Based Structural Equation Modeling ●

Section on Statistics and Marketing

Thursday, August 7, 8:30 a.m.–10:20 a.m.

VisualGSCA: A Graphical User Interface Software Program for Generalized Structured Component Analysis

✱Heungsun Hwang, McGill University, *heungsun.hwang@mcgill.ca*

Generalized Structured Component Analysis (GSCA) (Hwang & Takane, 2004) represents a component-based approach to structural equation modeling in which latent variables are defined as weighted composites of observed variables. GSCA is now implemented into a free software program called Visual GSCA 1.0. This program provides a graphical user interface whereby path analytic models are easily drawn in the program window and their GSCA parameter estimates are subsequently displayed in the same window. A brief introduction to GSCA is first presented. Next, the use of VisualGSCA 1.0 is demonstrated step by step through an empirical application.

Component-Based Structural Equation Modeling for Small Samples: A Comparison Between PLS, GSCA, and ULS-SEM

✱Michel Tenenhaus, HEC Paris, 1 rue de la Libération, Jouy-en-Josas, 78351 France, *tenenhaus@hec.fr*

Key Words: Structural Equation Modelling, PLS path modelling, Multiblock analysis

Two competing schools have come to the fore in the field of Structural Equation Modeling (SEM): covariance-based SEM and component-based SEM. The first approach has been developed around Karl Jöreskog and the second one around Herman Wold under the name “PLS” (Partial Least Squares). Hwang and Takane have proposed a new component-based SEM method named Generalized Structural Component Analysis. Covariance-based SEM is usually used with an objective of model validation and needs a large sample. Component-based SEM is mainly used for score computation and can be carried out on very small samples. In this research, we will explore the use of ULS-SEM, PLS, and GSCA on small samples. First experiences have shown that score computation and bootstrap validation are very insensitive to the choice of the method. We will also study the contribution of these methods for multiblock analysis.

The Contribution of PLS Regression to PLS Path Modelling: Formative Measurement Model and Causality Network in the Structural Model

✱Vincenzo Esposito Vinzi, ESSEC Business School of Paris and Singapore, Avenue Bernard Hirsch - B.P. 50105, Cergy-Pontoise, 95021 France, *vinzi@essec.fr*

Key Words: PLS path modelling, PLS regression, formative or emergent constructs, multidimensional blocks, multicollinearity

PLS Path Modeling (PLS-PM) is generally meant as a component-based approach to structural equation modeling that privileges a prediction oriented discovery process to the statistical testing of causal hypotheses. In case of formative relationships between manifest and latent variables, PLS-PM implies multiple OLS regressions. They might yield unstable results in case of strong correlations between manifest variables while being not feasible when the number of observations is smaller than the number of variables nor in case of missing data. We explore PLS regression (PLS-R) as an external estimation mode to overcome the mentioned problems while preserving formative relationships and being coherent with the component-based and prediction-oriented nature of PLS-PM. PLS-R is also fruitfully extended to the path weighting internal estimation scheme and the estimation of path coefficients.

Searching for a Group Structure in PLS Structural Equation Models: PLS Typological Path Modeling

✱Silvia Squillacciotti, EDF R&D, 1 avenue du Général de Gaulle, Clamart, International 92000 France, *silvia.squillacciotti@edf.fr*

Key Words: Partial Least Squares (PLS), Structural Equation Models (SEM), Typological modelling, Model-based classification, local models, customer satisfaction

PLS Structural Equation Modeling (PLS-SEM) has witnessed in recent years a growing number of applications in marketing, namely in the analysis of customer satisfaction and loyalty. The underlying assumption of PLS-SEM, that a unique model can correctly describe the behavior of all units, is often unrealistic. The failure to account for heterogeneity may lead to ambiguous results and wrong interpretations. Traditional techniques may be ineffective since unable to account for heterogeneity in the links between latent variables. In this work a method aiming at uncovering groups on the basis of the underlying model is described: PLS Typological Path Modeling (PLS-TPM). PLS-TPM leads to groups with distinctive PLS path models. An application to customer satisfaction modeling will be shown to illustrate the interest of the method.

The PLS-CBSEM Debate and an Investigation of Alternative Modeling Methods

✱Irene R.R. Lu, York University, 4700 Keele Street, Toronto, ON M1V 4J1 Canada, *irenelu@yorku.ca*; D. Roland Thomas, Sprott School of Business, Carleton University; Ernest Kwan, York University; Marzena Cedzynski, Sprott School of Business, Carleton University

Key Words: Partial Least Squares, Structural Equation Modeling, Latent Variable Modeling, Path Modeling, PLS, SEM

Partial least squares (PLS) is sometimes used as an alternative to covariance-based structural equation modeling (CBSEM). We identified and examined advantages of PLS over CBSEM commonly cited by PLS users from a broad range of social and behavioral disciplines. We further compared the performance of PLS and CBSEM to other latent variable modeling methods through a series of Monte Carlo simulations. The performance criteria comprised parameter estimation bias, standard error of parameter estimates, confidence interval coverage, and statistical power. Based on the simulation results we provide some recommendations on choosing methods for latent variable modeling.

472 Assessing Bias in Pre-Clinical and Clinical Diagnostic Studies ●▲

Section on Statistics in Epidemiology, Biopharmaceutical Section, WNAR, Biometrics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Different Schemes of Verification Bias in Evaluating Medical Tests

✱Marina V. Kondratovich, U.S. Food and Drug Administration, HFZ-550, 1350 Piccard Dr., Rockville, MD 20850, *Marina.Kondratovich@fda.hhs.gov*

Key Words: diagnostic accuracy, verification bias, sensitivity, specificity, multiple imputation, comparison of two tests

We discuss two types of verification biases in evaluating medical tests: partial verification bias and differential verification bias. With regard to the partial verification bias, we present different schemes of verification bias for a comparison of two medical tests (Old Test and New Test): 1) a random sample of the subjects with negative results via both tests has gold standard results; 2) no subjects with negative results via both tests have gold standard results; we describe two schemes: one for a paired study design (pre-market studies) and another for an unpaired study design (post-market studies); 3) only subjects with positive results by the Old Test have results of the gold standard. The results of investigation of conditions for the unbiased comparison of the two tests are presented. The differential verification bias is discussed with an example of prognostic markers.

Outliers in Method Comparison Studies

✱Geraldine Rauch, Roche Diagnostics GmbH, Biostatistics, Nonnenwald 2, Penzberg, International D-82377 Germany, *geraldine.rauch@roche.com*; Andrea Geistanger, Roche Diagnostics GmbH; Jurgen Timm, University of Bremen

Key Words: Outliers, Method Comparison, Local Variance Estimation

Evaluations of method comparison studies may be influenced by the presence of outlying measurements. This may result in biased precision or accuracy estimates. In order to identify outlying observations in advance, rules for outlier identification are needed. Method comparison studies are often evaluated by robust linear regression methods. Identification of outliers can be based on the construction of confidence intervals for the orthogonal residuals. These confidence intervals depend on the local residual variance, which is often difficult to estimate due to inhomogeneous sample distribution and heteroscedastic measurement variance. Techniques to calculate valid local residual variance estimates will be presented. Their utility is demonstrated by simulations and practical examples.

Partial Least Squares Observer for Detection of Breast Cancer Lesions in Images

✱Laura Thompson, U.S. Food and Drug Administration, HFZ-550, CDRH, 1350 Piccard Dr., Rockville, MD 20850, *laura.thompson@fda.hhs.gov*

Key Words: partial least squares, signal detection, imaging

We introduce a partial least squares (PLS) approach to estimate the Hotelling observer (classifier) for signal detection in medical imaging. The PLS observer selects orthogonal dimensions that maximize signal detection (maximize the covariance between the response and derived predictors). An advantage of the PLS observer over the commonly used channelized Hotelling observer (CHO) is that the selection of dimensions is done automatically, without user input. We will compare the performance of the PLS observer to the CHO and to the channelized ideal observer, in the detection of breast cancer in medical imaging.

Assessment of Surrogate Endpoint for Prediction of Clinical Outcome in Drug-Eluting Stent Trials

✱Hong Wang, Boston Scientific Corporation, 100 Boston Scientific Way, Marlborough, MA 01752, *Hong.Wang@bsci.com*; Don Wang, Boston Scientific Corporation; Carola Alfaro, Boston Scientific Corporation

Key Words: Logistic Regression, Simulation, Drug Eluting Stent, Surrogate Endpoint

In this paper, we use the proportion of treatment effect (PTE), likelihood reduction factor (LRF), proportion of information gain (PIG), receiver operating characteristic (ROC) curve, and the area under the ROC curve to assess the correlations of the quantitative coronary angiography (QCA) parameters with the clinical outcome of target lesion revascularization (TLR) in drug eluting stent trials. The examined QCA surrogate endpoints will include the analysis segment and stented region of late loss, percent diameter stenosis, and minimum lumen diameter. The PTE, LRF, PIG and c-statistic will be computed with its 95% confidence interval for each QCA parameter. Simulation will be applied to compare the performance of PTE, LRF and PIG. A threshold from QCA parameters to predict clinical outcome using the probability of misclassification minimization criterion will be explored.

Exploration of Different Test Methods on Ordinal Data from Preclinical Studies

✱Yongyi A. Yu, Boston Scientific Corporation, 15620 76th Pl N, Maple Grove, MN 55311, *yua@bsci.com*; So Jung Imm, Boston Scientific Corporation

Key Words: Ordinal Data, Frequency Table, Mantel-Haenszel Chi-square, Nonparametric method, Semi-nonparametric method

Preclinical studies on medical devices are often consisted of ordinal data with small sample sizes. Several analysis methods are explored and compared. Data from real world studies as well as from simulations are utilized to investigate the assumptions and validities of different methodologies.

473 Statisticians: Speaking Out and Reaching Out on Global Health

Section on Health Policy Statistics, Section on Statistics in Epidemiology, Social Statistics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Statisticians: Speaking Out and Reaching Out on Global Health Policy

✱Roger W. Hoerl, GE Global Research, 1 Research Circle, Niskayuna, NY 12309, *hoerl@crd.ge.com*; ✱Donna Stroup, Data for Solutions, Inc., *donnafstroup@comcast.net*; ✱Stephen Pierson, The American Statistical Association, *pierson@amstat.org*

Key Words: Health Policy, AIDS, Obesity, Tobaco, HIV

The purpose of this session is to bring together statisticians interested in global health care issues, such as HIV/AIDS, obesity, and chronic disease, to share experiences and discuss opportunities for statisticians to get more involved in reaching out beyond our profession, with the ultimate goal of having greater impact on public policy related to global health crises. The session will leave significant time for open floor discussion. Roger Hoerl (GE R&D) will share his experiences on a month-long journey to Africa to study HIV/AIDS - part of a 6-month sabbatical focused on HIV, and Donna Stroup (Data for Solutions) will discuss her experiences in the Caribbean and Eastern Europe, researching the consequences of obesity and smoking in the population. Stephen Pierson, ASA's Director of Science Policy, will serve as a discussant, while ASA President-elect Sally Morton will moderate.

474 Measurement Error Models and Matrix Sampling

Section on Survey Research Methods, Section on Government Statistics

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Evaluation of Error Components in a Simulation-Based Evaluation of a Survey Procedure

✱Moon J. Cho, Bureau of Labor Statistics, 9676 Scotch Haven Dr, Vienna, VA 22181, *Cho.Moon@bls.gov*; John L. Eltinge, Bureau of Labor Statistics

Key Words: Approximation error, Estimation error, International Price Program

In the evaluation of a survey procedure, simulation work rarely incorporates all of the features of the full superpopulation model, the true sample design, and the true estimation method. For example, practical constraints may limit a simulation study to cases in which finite populations are generated through independent and identically distributed realizations of a superpopulation model. Similarly, it may not be entirely feasible to replicate fully the design features. In addition, estimation procedures often involve complex computations of weights and other use of auxiliary information. Consequently, a given simulation study may not fully capture the sources of variability. We explore the extent to which the simulation-based distribution function provides an adequate approximation to the true distribution functions using selected data from the International Price Program.

A Protocol Calibration Experiment in a Longitudinal Survey with Errors-in-Variables

✱Cindy Yu, Iowa State University, Department of Statistics, 216A Snedecor Hall, Ames, IA 50010, *cindyyu@iastate.edu*; Jason C. Legg, Iowa State University

Key Words: Measurement error, Area sampling, Generalized least squares, National Resources Inventory

The National Resources Inventory (NRI) is a large-scale longitudinal survey conducted to assess trends and conditions of nonfederal land. A key NRI estimate is year-to-year change in acres of developed land, where developed land includes roads and urban areas. Since 2003, a digital data collection procedure has been implemented. Data from an NRI calibration experiment are used to estimate the relationship between data collected under the old and new protocols. A measurement error model is postulated for the relationship, where duplicate measurements are used to estimate the error variances. Analyses suggest that therefore the parameters in current use are acceptable. The paper also provides a way to model the measurement error variances as functions of the proportion of developed land, which is essential for estimating the effect of measurement.

Measurement Error Models for Physical Activity Assessments with Application to NHANES

✱Nicholas Beyler, Iowa State University, Ames, IA 50011, *beylern@iastate.edu*; Sarah Nusser, Center for Survey Statistics and Methodology, Iowa State University; Gregory Welk, Iowa State University

Key Words: Measurement Error, Physical Activity, NHANES

Physical activity (PA) data are important for studying health-related problems. Of interest is the long-run average of measures such as energy expenditure (EE) and time spent in moderate to vigorous physical activity (MVPA). Daily or weekly PA data are often obtained through recall methods in surveys, but these methods measure long-run PA with considerable error. Devices such as accelerometers yield data with less error, but are expensive to administer in a study. We consider measurement error models for recall and accelerometer-based data as a function of the unobservable long-run average of daily PA. We use data from the 2003-2004 National Health and Nutrition Examination Survey (NHANES).

An Application of Matrix Sampling and Multiple Imputation: The Decisions Survey

Mick P. Couper, The University of Michigan; ✱Trivellore Raghunathan, The University of Michigan, M4071 SPHII, 1420 Washington Heights, Ann Arbor, MI 48109, *teraghu@umich.edu*; John Van Hoewyk, The University of Michigan; Sonja Ziniel, The University of Michigan

Key Words: split questionnaire, matrix sampling, multiple imputation, planned missingness

We describe a split-questionnaire survey (Raghunathan and Grizzle, 1995) in a national RDD telephone survey of medical decision making among persons age 40+ in the US. Respondents who were eligible for up to 9 different medical conditions (3 cancer screenings, 3 medication, and 3 surgeries) were only asked detailed questions on up to 2 conditions. The selection of modules was inversely proportional to the expected marginal prevalence rate for each condition, and every pair of modules was administered across sampled individuals. Of the 3010 respondents, 1279 were eligible for 3 or more modules and were subject to random assignment. A total of 4528 of the 6696 possible modules were administered. We describe the selection algorithm and the subsequent multiple imputation to recover complete module-level information for all eligible modules on all 3010 respondents.

Multiple Imputation for Latent Variables in Classical Test Theory for Cluster Samples: A Simulation Study

✱Tiandong Li, Westat, 1650 Research Blvd., Rockville, MD 20850, *TiandongLi@Westat.com*; Robert Mislevy, University of Maryland

Key Words: Classical test theory, Cluster sample, Multiple imputation, Random effect model, Large scale assessment

Large scale assessments, such as NAEP, use complex sample designs, which classical measurement theory did not initially reflect. Mislevy (Psychometrika 1991) briefly introduced a random effects model for a latent construct under complex sample designs by applying multiple imputation techniques. An imputation statistic on ? was displayed without further discussion. For 2007 JSM, we demonstrated an analytical solution for the mean and variance of this imputation of ? for clustered samples, following Rubin's variance formula for multiple imputation. This study will extend this discussion under the classical test theory framework. With a simulation study, we will explore the effect on the variance of the imputed values by changing the relative sizes of the three variance components: between cluster variance, within cluster variance, and error variance.

Adaptive Matrix Sampling for the Consumer Expenditure Interview Survey

✱Jeffrey M. Gonzalez, Bureau of Labor Statistics, 2 Massachusetts Avenue NE, Suite 1950, Washington, DC 20212, *gonzalez.jeffrey@bls.gov*; John L. Eltinge, Bureau of Labor Statistics

Key Words: Adaptive Design, Burden Reduction, Multiple Imputation, Sample Survey, Two-Phase Sampling, Variance Estimation

The Consumer Expenditure Interview Survey is an ongoing panel survey of U.S. households in which sample units typically receive the same survey protocol during each interview. Because of the high burden associated with the survey request, the BLS is exploring alternative designs that, if implemented, would change many features of the data collection process. One such alternative is adaptive matrix sampling. Matrix sampling involves dividing a survey into sections and then administering each to subsamples of the main sample. To potentially compensate for the resulting loss of information, we can adapt section assignment probabilities based on data from the first interview. We use historical data to explore potential efficiency gains incurred by the use of adaptive matrix sampling, develop point estimators for expenditures collected under this design and evaluate their variance properties.

Imputing Responses to Nonexistent Questions

✱Peter Frechtel, RTI International, 701 13th Street NW, Suite 750, Washington, DC 20005, *frechtel@rti.org*

Key Words: imputation, model-based imputation, analysis of survey data

In surveys that are conducted annually, new questions may be added and old questions may be eliminated from year to year. For the sake of continuity and/or trend analysis, it may be helpful and informative to predict responses to these questions in the survey years when the question did not appear. In the National Survey on Drug Use and Health (NSDUH), an annual national survey of substance use and mental health measures in the U.S. civilian population aged 12 or older, a method has been used in which regression models are fit using data from the years when the question was included in the questionnaire, and predicted values are calculated for respondents during the years when the question was not included. The assumptions required for this method will be delineated and tested using real data from the NSDUH; and the analytic risks involved in such an approach will be noted.

475 Applications of Epidemiologic Models ●

Section on Statistics in Epidemiology, Biometrics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Predicting Access Rates to Substance Abuse Programs in Correctional Facilities Using Regression Trees

✱Panagiota Kitsantas, George Mason University, 4400 University Drive, College of Health and Human Services, Fairfax, VA 22030, *pkitsant@gmu.edu*; Faye Taxman, George Mason University; Matthew Perdoni, George Mason University; Deanna Breslin, George Mason University

Key Words: regression trees, substance abuse, correctional facilities, access rates

The purpose of this study was to investigate the relative importance of structural, administrative, and program level variables in predicting access rates to substance abuse treatment services for drug-involved offenders. The data, which were collected through a nationally representative survey of correctional and treatment administrators as part of the Criminal Justice Drug Abuse Treatment Studies, included 295 substance abuse treatment programs that were classified into three modalities of treatment services: high, medium and low intensity. Regression trees were built for each of these modalities to determine which variables were important in predicting access. This is the first phase of understanding capacity of correctional agencies to appropriate treatment services for offenders.

Cross-Validation of the Patient Health Questionnaire in Patients with Coronary Heart Disease

✱Yating Yeh, Statistics Lifesaver, Brookline, MA 02111, *statisticslifesaver@gmail.com*

Key Words: Patient Health Questionnaire, cross-validation, Confirmatory factor analysis, psychometric, Behavioral Risk Factor Surveillance System, exploratory factor analysis

We evaluated the psychometric properties of the Patient Health Questionnaire (PHQ), an instrument measuring the depressive symptoms, in Behavioral Risk Factor Surveillance System. The participants of the study were Washington residents who were diagnosed with coronary heart disease. The split-half, cross-validation method was applied. First, exploratory factor analysis was conducted to determine the measurement structure with the first half of the sample. A one-component solution was obtained. Cronbach coefficient Alpha was 0.81,suggesting good internal consistency. Second, the researcher used data from another half of the sample and applied confirmatory factor analysis to cross-validate the measurement structure. The one-factor model fitted the data well,?2(20)=68.72,CFI=0.91,GFI=0.93,RMSEA= 0.09. Results provided the evidence of the reliability and validity of PHQ.

A Variable Selection Macro for Logistic Regression in the SAS-Callable SUDAAN RLOGISTIC Procedure

✱Ruben Smith, Centers for Disease Control and Prevention, 4770 Buford HW, NE., Mail Stop K-21, Atlanta, GA 30341, *eyb4@cdc.gov*

Key Words: logistic regression, sample survey data, variable selection method

Automated variable selection methods are often used for choosing a subset from a large pool of explanatory variables in best model determination. Although for the past several years there has been variable selection software available for standard logistic regression, this capability is not available for the survey analyst due to the fact that standard logistic regression methods are inappropriate for the analysis of data arising from a complex sample design. This presentation will describe a macro that was written using the SAS-Callable SUDAAN RLOGISTIC procedure (which fits logistic regression models to sample survey data) to perform the variable selection algorithm described by Hosmer and Lemeshow (2000). The macro provides a summary at each step of the algorithm, and the best main effects logistic regression model is presented at the end of the output.

Alcohol Consumption, BMI, and Colorectal Cancer

✱Negasi T. Beyene, Centers for Disease Control and Prevention, P.O.Box 1696, Hyattsville, MD 20782, *nbeyene@cdc.gov*

Key Words: Colorectal Cancer, Alcohol consumption, BMI

Colorectal cancer is the third most common cancer and third leading cause of cancer-related mortality in the United States. Over the past decade, colorectal cancer incidence and mortality rates have modestly decreased. Until age 50, men and women have similar incidence and mortality rates; after age 50, men are more vulnerable. Different studies conclude differently if weight, height, and alcohol consumption are colorectal cancer risk factors. Using the National Health Interview Survey (NHIS) 1997-2006 we will examine the rate of colorectal cancer and if BMI and alcohol have anything to do with having colorectal cancer or not. We also examine if there is difference incidence of colorectal cancer among different races/ethnicities.

Multivariate Regression Analysis of Longitudinal Repeated Measures Data Modeling Relationships Among Weight Loss, Inflammatory Markers, and Type-2 Diabetes Mellitus

✱Ye-Ying Cen, Hennepin County Medical Center, 701 Park Ave, Minneapolis, MN 55404, *ye-ying.cen@hcmed.org*

Key Words: multivariate regression, mixed model, longitudinal repeated measures

In a randomized study of Type-2 diabetes mellitus, we hypothesized: 1) Weight loss is associated with reduction in markers of vascular inflammation (CRP, IL6), and 2) this relationship is influenced by insulin sensitivity and glycemic control (A1c). Mixed model for longitudinal repeated measures data was used. Month effect (baseline, 2, 6, and 12 months) is expressed through within subject variation. Two-response multivariate model was run and the results were compared to that of separate models. Residual plots and influential analysis were provided for model validation. Bivariate regression model gave similar results as two separate models for CRP and IL6. Weight loss is significantly associated with reduction of CRP, but not with IL6. The amount of reduction in CRP, however, depends on a1c. The higher the a1c is, the less the reduction will be.

Implications of Subdividing Marital Status: Are Unmarried Mothers with Partners Different from Unmarried Mothers Without Partners?

✱Robin Young, Boston University School of Public Health, 715 Albany St, Crosstown Center - 3rd Floor, 346c, Boston, MA 02118, *ryoung@bu.edu*; Eugene Declercq, Boston University School of Public Health

Key Words: Marital Status, Adverse Birth Outcomes, Unmarried Mothers

This study examined the analytic benefits of separating unmarried mothers with and without partners when assessing risk of adverse birth outcomes. Data from Listening to Mothers II, a nationally representative survey of mothers' childbearing experiences categorized marital status as married(71%), unmarried with partner(24%), or unmarried without partner(5%). Unmarried mothers with and without partners were similar in age, education, and parity but differed in method of delivery. Multiple logistic regressions controlling for demographics showed, compared to married mothers, unmarried mothers with partners had increased risk of giving birth to premature or low birth weight infants. Mothers without partners had even greater risk. Among primiparas, unmarried mothers with partners had twice the risk of premature infants and those without partners had 5 times greater risk than married mothers.

Causal-Effect Modeling for the Effects of Diabetes Oral Agents on Cardiovascular Diseases

✱Chen-Pin Wang, UTHSCSA, San Antonio, TX 78229, *cwang@verdict.uthscsa.edu*

Key Words: causal effects, counterfactual, diabetes

This presentation explores three causal-effect models to compare the effects of diabetes oral agents on cardiovascular diseases directly and indirectly via glycemic control in a clinical cohort with type 2 diabetes. All three models considered are developed from the counterfactual framework. However, they differ in how the counterfactual outcomes are estimated, and how the sample should be stratified and then matched based on treatment assignments as well as the intermediate outcome: glycemic control. We demonstrate the impact of model misspecification on the estimates of direct and indirect treatment effects.

476 Environmental Time Series

Section on Statistics and the Environment
Thursday, August 7, 8:30 a.m.–10:20 a.m.

Temporally Correlated Dirichlet Processes for Pollution Receptor Models

✱Matthew Heaton, Duke University, 5012 Stardust Drive, Durham, NC 27712, *matt@stat.duke.edu*

Key Words: Dirichlet process, Receptor model, Autocorrelation, Air pollution, Dynamic model

In order to correctly model air pollution data, the temporal structure of pollution concentration measurements needs to be accounted for. This article extends the multivariate pollution receptor model by accounting for the temporal structure of air pollution data using a time-dependent Dirichlet process to model pollution source profiles. The Dirichlet process model developed herein is first evaluated using simulated data sets and is then applied to a physical data set of measurements of particulate matter collected near St. Louis. The Dirichlet process model is found to be preferable to more

traditional receptor models because it requires fewer assumptions for its use, more accurately estimates pollution source contributions and profiles, and is flexible enough to estimate complex quantities through the employment of Markov chain Monte Carlo parameter estimation techniques.

A Model-Based Approach for Clustering Time Series of Counts

✱Sarah J. Thomas, Rice University, Dept of Statistics MS-138, PO Box 1892, Houston, TX 77251-1892, *sarah.thomas@rice.edu*; Bonnie Ray, IBM T. J. Watson Research Center; Katherine B. Ensor, Rice University

Key Words: count time series, model-based clustering, observation-driven Poisson, zero-inflated Poisson

Time series of counts arise in many applications, such as modeling product purchases for a customer or tracking abundance of biological species. It is often important to identify those series exhibiting similar behavior over time. We model the count series with an observation-driven Poisson regression model, which incorporates the autoregressive component of the series into the mean process of the Poisson. The time series are then clustered with a hierarchical amalgamative clustering algorithm using an empirical Kullback-Leibler distance metric. The fit of the data to a given model is calculated with a partial likelihood technique. The possibility of fitting and clustering zero-inflated models is also explored. A goodness-of-fit test for the zero-inflated models is discussed. The algorithm is applied to modeling and clustering air pollution data for the Houston metropolitan area.

Modeling Population Trends by Combining Information from Multiple Time Series

✱Philip Dixon, Iowa State University, Snedecor Hall, Ames, IA 50011, *pdixon@iastate.edu*; David Otis, Iowa State University

Key Words: state space model, detection probability, combining indices

In the United States, two long-term, nationwide surveys provide three indices of Mourning Dove abundance. These are the Call Count Survey, which provides an index of birds heard and an index of birds seen, and the Breeding Bird Survey. Since 2003, these have been augmented by a direct estimate of population size from band return and harvest information. All four time series are imprecise; the three long-term population indices are potentially biased. We describe a state-space model to integrate all four sources of information into an estimated population trend. The contribution of a time series is down-weighted if that series is estimated to have a large bias or large variance. An extension to the model allows detection probabilities to vary over time. The performance and characteristics of the model are illustrated using data from the U.S. Western Management Unit.

Time-Frequency Analysis by Multiscale Methods with Applications to Bat Signals

✱Minjeong Park, Seoul National University, Gwanak-gu Shinlim-dong San 56-1 24-206, Department of Statistics, Seoul, 151-747 South Korea, *pmj1210@snu.ac.kr*; Sinsup Cho, Seoul National University; Heeseok Oh, Seoul National University

Key Words: instantaneous frequency, wavelet packets, cosine packets, empirical mode decomposition, Hilbert transform

A better estimation of the instantaneous frequency (IF) and time-varying amplitude of time series is very crucial for analyzing nonstationary signals, such as bat signals. Previously, various methods have been developed to estimate IF and amplitude. However, many existing methods focus on a simple visual display of the data on time-frequency plane, which might not be appropriate to understand the localized frequency and amplitude. In this paper, we propose a new statistical approach to estimate time-varying IF and amplitude based on multiscale methods coupled with parametric and nonparametric models. The empirical performance of the proposed method is tested on numerical studies and real bat signals. Furthermore, a new approach of best basis selection is addressed.

Evaluating Estimators of Influenza-Associated Mortality via Simulation

✱Jian Rong, Boston University School of Public Health, 715 Albany Street, Crosstown Center, Biostatistics, 3rd floor, Boston, MA 02118, *jrong@bu.edu*; Xiaopeng Miao, Boston University School of Public Health; Al Ozonoff, Boston University School of Public Health

Key Words: Influenza, Surveillance, Excess mortality, Simulations, Evaluation

Statistical methods for influenza surveillance, for example estimators of excess mortality, can be challenging to evaluate since the true parameter value of interest is unknown. We simulate weekly influenza mortality data, where the true value of excess mortality is specified by design, by generating time series with dynamic characteristics similar to historical data from the 122 Cities Mortality Reporting System (CMRS) operated by the Centers for Disease Control and Prevention (CDC). We assume a basic sinusoidal model with sporadic epidemic periods that are faithful in timing, duration, and amplitude to the empirical data. We then apply this simulation scheme to study bias and variability of different robust regression procedures as they are implemented to estimate influenza-associated mortality.

Additive Mixed Models for Assessing Change in Yearly Streamflow Timing Measures in the Western United States

✱Mark C. Greenwood, Montana State University, Bozeman, Department of Mathematical Sciences, Montana State University, Bozeman, MT 59717-2400, *greenwood@math.montana.edu*; Johnnie Moore, University of Montana; Joel Harper, University of Montana

Key Words: hydrology, additive mixed models, spatial-temporal models, climate change

Recent analyses of regional snow and river data for the late 20th century have suggested that western North American snow pack has decreased (Mote et al., 2003) and the spring snow melt pulse has arrived earlier (Stewart et al., 2005), but the results were less conclusive when adjustments for multiple testing were made in Moore et al. (2007). Streamflow discharge timing data for snow-melt dominated locations are considered for 1951 to 2005 on 21 gages located on the Missouri and Columbia headwaters. Additive mixed model methods (Wood, 2006) are used to estimate potential change over this time period and account for the impacts of total yearly streamflow discharge and climatological forcing functions such as PDO and ENSO. Additionally, spatial-temporal correlation structures are considered.

Comparing Pollution Levels Monitored at Sites Located on the U.S.-Mexico Border

✱Herbert Lacayo, U.S. Environmental Protection Agency, US EPA, MC2842T, 1200Pennsylvania Ave., NW, Washington, DC 20460, *lacayo.pepi@epa.gov*; Thomas Forbes, U.S. Environmental Protection Agency; Nagaraj K. Neerchal, University of Maryland, Baltimore County; Barry D. Nussbaum, U.S. Environmental Protection Agency

Key Words: Ozone, time series, Border Environment

Ozone data from several sites along the US-Mexico border are analyzed using a time series approach. The analysis disaggregates prior data taken from 1999–2006, and using ANOVA investigates site and seasonal differences. Simultaneous confidence intervals for the site means are presented to facilitate ranking of the ozone sites.

477 Statistical Methods and Applications in Defense and National Security ●

Section on Statisticians in Defense and National Security

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Computationally Efficient Resource Allocation for Complex System Reliability Studies

✱Jessica Chapman, Iowa State University, Ames, IA 50011, *jessicac@iastate.edu*; Max Morris, Iowa State University; Christine Anderson-Cook, Los Alamos National Laboratory

Key Words: System Reliablity, Markov Chain Monte Carlo, Resource Allocation

We introduce a new approach for assessing the effect of adding new data to an existing analysis when estimating complex system reliability. It is important and challenging to allocate resources among different data sources so as to obtain the most information about the reliability of the system for the resources available. One challenge is that the current approach for evaluating candidate allocations is computationally intensive, involving repeated runs of an MCMC algorithm. We propose an approximation that uses existing information and/or data to estimate the effect that new data would have on the posterior distribution for the system reliability. Once the posterior has been updated for a candidate allocation, it can be summarized to determine which candidate allocation is most informative. The result is that candidate allocations can be evaluated via a single run of an MCMC algorithm.

Development of a Forecasting Methodology for Marine Corps Equipments Demands

✱Steven M. Rollins, Marine Corps Logistics Command, Commander L70, 814 Radford Blvd Suite 20301, Albany, GA 31704, *steven.m.rollins@usmc.mil*

Key Words: Regression, Forecasting

The need for forecasting of Marine Corps demands of equipment is essential to insure that needed supplies are available when they are required. Currently the Marine Corps Logistics Command utilizes a variation of a double exponential forecasting methodology which delivers an accuracy of 30%. A new forecasting methodology is being developed that has preliminary results of 70% accuracy. The new forecasting methodology was inspired by Dynamic Bayesian Forecasting. The methodology is currently being fully developed and will be completed by 30 June 2008.

Experiment Design and Sampling Plans for a Building Contamination and Decontamination

✱Brett Amidan, Pacific Northwest National Laboratory, Richland, WA 99352, *b.amidan@pnl.gov*; Greg F. Piepel, Pacific Northwest National Laboratory; Brett Matzke, Pacific Northwest National Laboratory

Key Words: Experimental design, Sampling plan, National security

An experimental design was developed to assess sampling strategies and methods for (i) detecting contamination in a building and (ii) clearing the building for use after decontamination. An unoccupied building was contaminated with a Bacillus anthracis simulant. The experimental design planned several contamination events, with sampling, decontamination, and re-sampling for each. The first objective was to assess the abilities of judgmental and probabilistic sampling strategies to detect contamination. The second objective was to assess the use of a probabilistic sampling strategy and a Bayesian sampling strategy (which combines judgmental and probabilistic samples) to make clearance statements of the form "X% confidence that at least Y% of an area does not contain detectable contamination." The presentation gives an overview of the experimental design and sampling plans selected.

Approximation of Data Modeled by Functions with Removable or Jump Discontinuities

✱Patricia H. Carter, Naval Surface Warfare Center, 18444 Frontage Rd, Suite 327, Dahlgren VA, 22448-5161, *patricia.h.carter@navy.mil*

Key Words: piecewise linear approximation, time series, compression, segmentation, discontinuous data, data reduction

There are many algorithms for the compression of time series data by approximation using piecewise linear functions. As usual the approximation trade off is compression rate versus a global measure of the error as well as the resulting errors in determining size and locations of events of interest. Here we introduce an agglomerative approximation algorithm designed for time series data that is well-modeled by functions with removable or jump discontinuities. This algorithm is exact on appropriate piecewise linear models with isolated removable and jump discontinuities, and outperforms generic bottom-up and random iteration methods at similar computation cost and algorithmic complexity.

An Efficient Statistical Technique for Automated Band Detection in Remotely Sensed Imagery

✱Pranab K. Banerjee, Space Dynamics Laboratory, 1695 North Research Park Way, North Logan, UT 84341, *Pranab.Banerjee@sdl.usu.edu*

Key Words: Statistical image processing, band detection, remote sensing, pattern recognition, sensor calibration

Banding in remotely sensed imagery refers to the appearance of brighter or darker regions of fixed width spanning an entire image horizontally or vertically. These artifacts are most common in imagery from push-broom class sensors which are common in space-borne imaging platforms. If any segment of such a sensor suffers from sensitivity/calibration errors, the corresponding pixels will produce banding, and hence their existence is indicative of sensor errors. Because of the enormity of imagery collected by modern sensors, automated detection of these artifacts is essential for assessing sensor health. This paper presents an efficient statistical algorithm for detecting such bands. It requires no a-priori knowledge of their widths or numbers. It is based on single link agglomerative hierarchical clustering in a feature space based on ternary valued "pixel transition vectors."

Automated Recognition of Pedestrian Traffic from an Unmanned Ground Vehicle

✱Barry A. Bodt, U.S. Army Research Laboratory, ATTN: AMSRD-ARL-CI-CT, APG, MD 21005, *barry.a.bodt@us.army.mil*

Key Words: robotics, experimental design

An important consideration in the use of any unmanned ground vehicle is safety for those in proximity to the vehicle. Pedestrian tracking and avoidance is pursued using a variety of sensors and algorithms to process the sensor returns. For military application, the U.S. Army Research Laboratory in conjunction with industry and academic partners has conducted an extensive experimental investigation into the performance of various approaches. In this talk I will describe the design and analysis for this investigation.

478 Graphics for Regression, Classification, and Dimension Reduction

Section on Statistical Graphics, Section on Nonparametric Statistics

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Location-Dispersion Plots in Quadratic Classification Problems

✱Santiago Velilla, Universidad Carlos III De Madrid, C/Madrid, 126, 28903 Getafe Madrid, Getafe, International 28903 Spain, *santiago.velilla@uc3m.es*

Key Words: Data representation, Dimension reduction, Orthogonality, SAVE, SIR, SIRII

The structure of the quadratic subspace in discriminant analysis is studied from different perspectives, and related to several subspaces that have been proposed previously in the literature. A new decomposition of the quadratic subspace is introduced, and its usefulness for producing in practice interpretable location-dispersion plots considered. Special attention is given to the analysis of the standard Fisher's Iris Data Set.

Small Sample Histogram Possibilities and Paradoxes

✱James Weber, University of Illinois, PO Box 603, Gurnee, IL 60031-0603, *jsweber@uic.edu*

Key Words: Histogram variability, bin frequency level sets, histogram density estimation, small samples, symmetry, paradox

Small sample histograms may have inconsistent bin frequencies, due to variability and sensitivity to bin width and location. This might not be fully appreciated. Histogram variability and inconsistencies may be studied from an complete list of possible bin frequencies constructed via the bin frequency level set partition of the two dimensional space of bin locations and widths, {(t0,h) | h > 0}. Paradoxes include appearance skewness reversals, sometimes due only to a change of bin location. Graphic appearance error in terms of non-representative paradoxical histograms is discussed. (A byproduct of the level sets is a means of exactly determining the unique bin location and width values that minimize histogram Unbiased Cross Validation ("UCV") Mean Integrated Squared Error ("MISE").) Implications of symmetry suggest that normal populations could be vulnerable to histogram paradoxes.

Resultant-Vector Banking of Graphical Displays: Geometry and Statistical Properties

✱Saptarshi Guha, Purdue University, Dept. Of Statistics, 250 N. University Street, West Lafayette, IN 47907, *sguha@purdue.edu*; William Cleveland, Purdue University

Key Words: banking to 45 degrees, graphical displays, aspect ratio, graphical perception

The aspect ratio of a two-dimensional graph is critical when slopes are judged to determine how one variable depends on another, carried out by assessing the orientations of line segments. Theory and experiment have demonstrated that judgments are enhanced by banking to 45 degrees: the aspect ratio is chosen so that the absolute values of the orientations are centered on 45 degrees. A new approach to centering has a geometric motivation. Take the segments on the graph with any aspect ratio and think of them as vectors at an origin pointing into Quadrant I with polar angles equal to the absolute values of their orientations. Choose the aspect ratio on this construction so that the resultant vector has an angle of 45 degrees. Resultant-vector banking provides an excellent centering, geometrical insight into banking, and simple formulas that allow tractable theoretical study.

Regression Graphics for Bias Reduction in Observational Studies

✱Siamak Noorbaloochi, VAMC and University of Minnesota, 2E-1-7, One vetran Drive, Minneapolis, MN , *Siamak@stat.umn.edu*; David Nelson, Minneapolis VA Medical Center

Key Words: Regression Graphics, bias reduction, dimension reduction, casual analysis, propensity analysis, treatment comparisons

Stratification, covariance analysis, matching and propensity analysis are some common methods of bias reduction when randomized balanced treatment assignment is absent. Propensity analysis is a conditionally specified bias reduction method in casual analysis. We consider the bias reduction in treatment comparisons when a large number of imbalance baseline predictors are available. Based on the marginal distribution for the covariates, dimension reduction methods in Regression Graphics will be used as an alternative to propensity analysis for reducing the bias in treatment comparisons. We show how sliced inverse regression (SIR) and other dimension reduction estimation methods such as SAVE, p.H.d and SIR-II for constructing effective directions, in addition to dimension reduction, can be used to reduce bias and, under ignorability condition, perform casual analysis.

The Central Limit Theorem and Structural Validity in Cluster Analysis

✱Ryan Shatzer, Brigham Young University, 1001 SWKT, Psychology Department, Provo, UT 84602, *ryanshatzer@hotmail.com*; Robert Bubb, Brigham Young University; Michael Lauritzen, Brigham Young University; Bruce L. Brown, Brigham Young University

Key Words: cluster analysis, structural integrity, reliability, Toeplitz structure, clustered structure

The central claim of cluster analysis is that it can identify the clustered or hierarchical structure of data from generalized Pythagorean distances among observed datapoints. Monte Carlo simulation is used to create surface observations from a known systematic underlying structure, either two-dimensional or three-dimensional, with the variables in clustered or Toeplitz structure, with one of six levels of measurement error added to the data. Quantitative and also graphical methods are used to assess the extent to which structural integrity is maintained for each type of data at each level of measurement error. The watershed value seems to be a reliability of .50. Below that level the structural integrity of the solution deteriorates rapidly.

Factorial Complexity and Structural Integrity as Reflected in Graphs of Eigenvalues and Graphs of Telescoping Determinants and Traces

✱Daniel S. Allen, Brigham Young University, 262 East 100 South #4B, Provo, UT 84606, *daniel.s.allen@gmail.com*; Dustin Fife, Brigham Young University; Robert Bubb, Brigham Young University; Bruce L. Brown, Brigham Young University

Key Words: factor analysis, eigenvalues, determinant of a matrix, trace of a matrix

Skree plots (graphs of successive eigenvalues) are often used as a method of determining how many factors should be extracted from a correlation matrix or a covariance matrix. However, the determinants and the traces of each diagonal minor of the eigenvalues matrix corresponding to a covariance or correlation matrix also have diagnostic information concerning factorial complexity and even structural integrity. Using Monte Carlo simulation, and data with various levels of structural integrity and reliability, convergent graphs of eigenvalues and telescoping determinant and trace indices are demonstrated as diagnostic tools.

479 Reform, Redesign, and Innovation in Teaching and Assessment of Learning ●

Section on Statistical Education

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Introductory Statistics: Perspectives of Departments, Instructors, and Students

✱Jennifer L. Green, University of Nebraska-Lincoln, 340 Hardin Hall, Lincoln, NE 68583, *jenwarriner@hotmail.com*; Walter Stroup, University of Nebraska-Lincoln; Erin Blankenship, University of Nebraska-Lincoln

Key Words: Introductory Statistics, Mixed Methods, Teaching Assistants, Survey, Focus Group

Recent reforms in statistics education and introductory statistics courses have initiated the need to prepare graduate teaching assistants (TAs) for these changes. Using a focus group study, we explored the experiences and perceptions of the University of Nebraska-Lincoln (UNL) TAs. Survey data from students taking a non-calculus introductory statistics course support the findings from this study, prompting a need for possible course changes. Factor analysis, mixed models, focus groups and traditional statistical methods were used to demonstrate the students' and TAs' support of a more concise, thorough approach toward technology, content and pedagogy in the introductory statistics course. Improvements are suggested to satisfy department, instructor, and student needs, as well as to promote critical thinking in statistics courses and create a statistically literate society.

Redesigning Entry-Level Statistics Courses Using Software

✱Kim Denley, University of Mississippi, Hume 217, University, MS 38677, *kdenley@olemiss.edu*

Key Words: redesign, statistics, computer software

"Redesign." The buzz word for the year. What does that look like on a big scale? How can you get more students to do more work for entry-level statistics courses?The University of Mississippi has been in the process of redesigning entry level mathematics courses for several years now. We've run trials of several different methods as well as efforts to maximize student outcomes and instructor resources. This presentation will showcase what has been learned and will seek to help others explore how to redesign their own introductory statistics courses.

Lancaster Postgraduate Statistics Centre: Creating Enterprise and Innovation in Teaching Statistics Across Disciplines

✱Gillian Lancaster, Lancaster University, Postgraduate Statistics Centre, Fylde College, Lancaster, LA1 4YF United Kingdom, *g.lancaster@lancaster.ac.uk*

Key Words: CETL, HEFCE, higher education, postgraduate statistics, skills acquisition

In 2002 the Higher Education Funding Council of England launched its largest ever single funding initiative in teaching and learning, amounting to £315 million to fund 74 new centers from 2005-2010. The Centres for Excellence in Teaching and Learning (CETL) initiative has two main aims: to reward excellent teaching practice, and to further invest in that practice to deliver substantial benefits. The Lancaster Postgraduate Statistics Centre is the only CETL that specializes in postgraduate statistical education and builds on the teaching excellence and well-established 6* RAE rated research reputation of the Statistics Group. The diverse interests of the group provide a unique research-led approach to teaching with strong practical emphasis on skills acquisition and promoting PG statistics across disciplines. This talk will describe the breadth and diversity of the work of the new centre.

A Modified Moore Approach to Teaching Mathematical Statistics

✱Padraig McLoughlin, Kutztown University of Pennsylvania, 116C Lytle Hall, Dept. of Mathematics, Kutztown University of Pennsylvania, Kutztown, PA 19530, *mcloughl@kutztown.edu*

Key Words: Education, Moore Method, Probabilty, Mathematical Statistics, Inquiry-based learning, student-centered learning

The author argues a student needs to learn to conjecture, hypothesize, and prove or disprove claims. Ergo, the paper's thesis is that learning requires doing; only through inquiry is learning achieved, and hence proposed is a program of use of a modified Moore method (MMM) in a two course Mathematical Statistics Sequence (MSS) to teach students how to do, critique, or analyze arguments, examples, or their denials in the foundations of probability and statistical theory. Pedagogical and practical justification are submitted; the model for the two course MSS is detailed (what is effective, why, & what practices were refined, modified, or deleted); and, the successes or lack thereof are discussed in the MSS courses. It is forwarded that the MMM establishes an atmosphere that creates an easier transition to advanced statistics and assists in forging meaningful student research.

Proficiency-Based Assessment and Reassessment of Learning Objectives in an Introductory Statistics Class

✱Michael A. Posner, Villanova University, 800 Lancaster Ave, Villanova, PA 19085, *michael.posner@villanova.edu*

Key Words: Assessment, Statistics Education, Formative Assessment, Standards

This research study examines a proficiency-based grading system of learning objectives including the opportunity for assignment resubmission and compares it to a traditional single-attempt, numeric, summative assessment methodology in two introductory statistics classes. Results showed that students in the two classes performed equally well on the final exam, when controlling for past performance in math classes. Within the experimental class, students who learned the material only after a second submission learned it just as well as students who grasped the material the first time and students who chose to resubmit assignments performed better on the final exam. Student and instructor benefits from defining learning objectives were also observed. Preliminary results from the second year of data gathering will be included as well.

Using Artist CAOS Tests To Inform Classroom Instruction and Measure Student Improvement Over Two Semesters

✱Michelle Sisto, International University of Monaco, 2 Avenue Prince Albert II, Monte Carlo, 98000 Monaco, *msisto@monaco.edu*

Key Words: Assessment, Statistical Reasoning, ARTIST, On-line Tests

The push toward greater emphasis on statistical literacy, thinking and reasoning has taken off in the Statistics Education community in recent years.

The Assessment Resource Tools for Improving Statistical Thinking (ARTIST) website provides many easy to use online resources for instructors. Since January 2005 the International University of Monaco has been using the Comprehensive Assessment of Outcomes for a first course in Statistics (CAOS) pre and post tests for all students over our two-semester introductory statistics sequence. Results of these tests prove useful in measuring overall improvement of individual students as well as providing clear indications of concepts the class as a whole mastered or did not master. I shall present examples of how we use CAOS results to inform our teaching and will share initial findings of our students' performance over the last 3 years.

Readings in the Theory of Cognition for Teachers of Statistics

✱Gary Fowler, U.S. Naval Academy, Mathematics Department, 572 Holloway Road, Annapolis, MD 21402, *gof@usna.edu*

Key Words: cognition, learning, teaching

In the last few decades psychologists have made important progress in their understanding of the many mental activities related to cognition in general and the learning of statistics in particular. Many authors of modern statistics textbooks are well acquainted with this science. Frequently teachers are acquainted only with summaries of the science of cognition. In a sense this situation is similar to a scientist (statistics teacher) using a statistical software package (modern statistics textbook) with only a cursory understanding of statistics (cognition). This talk presents suggested readings in cognition for the statistics teachers with the goal of understanding the basic science. As it happens this literature contains research that make good case studies that can help both in the instruction of statistics and an understanding of what is entailed in the learning process.

480 EM Algorithms and Their Applications

Section on Statistical Computing, Section on Bayesian Statistical Science

Thursday, August 7, 8:30 a.m.–10:20 a.m.

A Robust Nonlinear Mixed-Effects Model Using the SAEM Algorithm

✱Cristian Meza, Universidad de Valparaíso, Departamento de Estadística, Gran Bretaña 1091, Valparaíso, 2340000 Chile, *Christian.meza@deuv.cl*; Rolando de la Cruz-Mesia, Pontifícia Universidad Católica de Chile; Felipe Osorio, Universidad de Valparaíso

Key Words: Nonlinear mixed models, SAEM algorithm, Robust model, Repeated measures, Longitudinal data, Outliers

The nonlinear mixed-effects models are very useful to analyze repeated measures data. Usually, it is assumed normal distributions for the random effects and the residuals, but such assumptions make inferences vulnerable to the presence of outliers. We introduce an extension of normal nonlinear mixed-effects model considering a subclass of elliptical contoured distributions for both random effects and residual errors. This elliptical subclass, the scale mixtures of normal (SMN) distributions, includes heavy tailed multivariate distributions, such as Student-t and contaminated normal among others. We propose an exact estimation procedure to obtain the maximum likelihood estimates of the fixed effects and variance components using a stochastic approximation of the EM algorithm. We compare the performance of the normal and the SMN models and it is shown our model allows to identify outliers.

On a Mixture of Skew T Distributions

✱Wan Ju Hsieh, National Chiao Tung University, Institute of Statistics, Hsinchu, 300 Taiwan, *siehcandy.st93g@nctu.edu.tw*; Tsung-I Lin, National Chung Hsing University

Key Words: EM-type algorithms, maximum likelihood, outlying observations, PX-EM algorithm, skew t mixtures, truncated normal

A finite mixture model using the Student's t distribution has been recognized as a robust extension of normal mixtures. Recently, a mixture of skew normal distributions has been found to be effective in the treatment of heterogeneous data involving asymmetric behaviors across subclasses. In this article, we propose a robust mixture framework based on the skew t distribution to efficiently deal with heavy-tailedness, extra skewness and multimodality in a wide range of settings. Statistical mixture modeling based on normal, Student's t and skew normal distributions can be viewed as special cases of the skew t mixture model. We present analytically simple EM-type algorithms for iteratively computing maximum likelihood estimates. The proposed methodology is illustrated by analyzing a real data example.

A New Estimation Method for the Accelerated Failure Time Frailty Model

✱Linzhi Xu, University of South Carolina, 800 Sumter Street Rm. 211C, Columbia, SC 29208, *xlzhorse@hotmail.com*

Key Words: Frailty, Accelerated failure time model, EM algorithm, Rank estimation

Frailty model has been extensively applied to analyze correlated or clustered failure time data. In this paper, we develop a new estimation method for the accelerated failure time frailty model based on the EM algorithm and the rank estimation method. The rank-like estimating equation in the M-step can be performed easily through the linear programming method. Simulation results show that the proposed method outperforms other existing estimation methods. For illustration, this method is subsequently applied to the data from a coronary heart disease study.

A Simplified Stochastic Approximation Algorithm for Estimation of GLMMs

✱Sheng-Mao Chang, National Cheng Kung University, Department of Statistics, NCKU, No. 1 Ta-Hsueh Road, Tainan, 70101 Taiwan, *smchang@stat.ncku.edu.tw*

Key Words: stochastic approximation, GLMM, geometric ergodic

Consider the optimization problem when the objective function can only be measured with error. Stochastic approximation (SA) algorithm is designed for this sort of problems. In our previous work, we define a stationary version of SA and show that when the objective function is quadratic the solution sequence is stationary. In this work, we define the simplified stationary SA (S3A) algorithm and investigate the resulting distance (between the current solution and the true solution) sequence. With some mild conditions, the distance sequence is Markovian and is geometric ergodic. This provides the error bound of the S3A solution as well as a relevant stopping criterion. We demonstrate S3A by maximizing the marginal likelihood of generalized linear mixed models.

Comparing Variants of the EM Algorithm for Finite Mixtures of Linear Mixed Effects Models

✱Bettina Grün, Wirtschaftsuniversität Wien, Department für Statistik und Mathematik, Augasse 2-6, Wien, 1090 Austria, *Bettina.Gruen@wu-wien.ac.at*

Key Words: finite mixture, EM algorithm, mixed-effect model

Finite mixtures of linear mixed models are increasingly applied in different areas such as marketing or bioinformatics. For Maximum Likelihood estimation different variants of the EM algorithm have been proposed which differ in the amount of missing data included and in the complexity of the E- and the M-step. We outline the implementation of two variants where a general framework for EM estimation of mixtures as well as functionality for fitting linear mixed effects models is available. The performance of the two variants is compared using different criteria such as the solutions detected and the number of iterations required for convergence.

Fast Ml Estimation for the Mixture of Factor Analyzers via an ECM Algorithm

✱Jian-hua Zhao, The University of Hong Kong, RM 518, Meng Wah Complex, Department of Statistics and Actuarial Science, Hong Kong, China, *jhzhao.ynu@gmail.com*; Philip L.H. Yu, The University of Hong Kong

Key Words: maximum likelihood estimation, MFA, EM, AECM, ECM

We propose a fast expectation conditional maximization algorithm (ECM) for maximum likelihood (ML) estimation of mixtures of factor analyzers (MFA). Unlike the existing EM algorithms such as the EM in (Ghahramani and Hinton, 1996) and the alternating ECM (AECM) in (McLachlan and Peel, 2003), where the missing data contains latent class labels as well as latent factors, the missing data in our ECM consists of latent class labels only. The novelty of our algorithm is that close-form expressions in all conditional maximization (CM) steps are obtained explicitly, instead of resorting to numerical optimization methods. As revealed by experiments, the convergence of our ECM is substantially faster than EM and AECM no matter assessed by CPU time and number of iterations.

481 Advances in Reliability and Life Testing ●

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Bayesian Analysis for Step-Stress Accelerated Life Testing Using Weibull Proportional Hazard Model

✱Naijun Sha, University of Texas at El Paso, *nsha@utep.edu*; Rong Pan, Arizona State University

Key Words: Step-stress accelerated life test, proportional hazard model, Bayesian inference method

In this paper, we present a Bayesian analysis for Weibull proportional hazard model in step-stress life testing. Compared with Weibull cumulative exposure model, this alternative model has attractive mathematical properties to be easily handled in the Bayesian framework. A MCMC algorithm with adaptive rejection sampling is used for posterior inference. We demonstrate the performance of the method on both simulated and real data sets.

Planning Accelerated Degradation Tests with Random Coefficient Models

✱Shuen-Lin Jeng, National Cheng Kung University, Dept. of Statistics, 701 Taiwan, *sljeng@mail.ncku.edu.tw*

Key Words: accelerated degradation tests, asymptotic variance, optimal test plans, random coefficient

Estimating the long term performance of highly reliable devices has been a difficult problem. The accelerated life tests, which involves testing at highly elevated stress, often results in too few failures. To overcome this problem, accelerated degradation tests (ADTs) take measurements along experiment to exhibit degradation and monitor it over time. These data provide valuable information on the failure mechanisms of the devices. In this study, we focus on the optimal test plans of accelerated degradation tests with random coefficient models. The position of the lower stress of a ADT design is obtained by minimizing the asymptotic variance of a particular percentile at the use-condition. The distribution of random coefficients is modeled by multivariate normal. Simulation study is conducted to show the design points under various values of the distribution parameters.

Weibull Prediction Bounds in Accelerated Life Testing with Two Stress Factors

✱Ananda A. Jayawardhana, Pittsburg State University, Dept. of Mathematics, Pittsburg State University, Pittsburg, KS 66762, *ananda@pittstate.edu*

Key Words: Prediction intervals, Accelerated life tests, Weibull distribution, Two-way experiemnts, Type II censoring

This method assumes a product life of a Weibull distribution and two factors of acceleration. Using the inverse power law model the scale parameter is modeled as a function of the stress factors. Using the maximum likelihood predictive density a predictive density is derived and used to find percentile points. Results of a simulation study will be presented. Probability coverage using percentile points of the predictive density are slightly liberal sometimes and therefore an ad-hoc method is proposed to improve the coverage very closer to the nominal levels.

The Beta-Rayleigh Distribution in Reliability Measure

✱Alfred Akinsete, Marshall University, Department of Mathematics, 1 John Marshall Drive, Huntington, WV 25755, *akinsete@marshall.edu*; Charles Lowe, Marshall University

Key Words: Beta distribution, Rayleigh distribution, Probability, Reliability, Moments, Parameter estimation

The problem of estimating the reliability of components is of utmost importance in many areas of research, for example in medicine, engineering and control systems. If X represents a random strength capable of withstanding a random amount of stress Y in a component, the quantity R = Prob(Y < X) measures the reliability of the component. Reliability measure has been discussed in the literature, with the random variables X, and Y having some known distributions. In the present paper, we define the beta-Rayleigh distribution (BRD), and obtain the measure of reliability with random variables having this new generalized distribution. Some properties of the BRD are also discussed, including for example, the mean, variance, moments, skewness, and kurtosis. The parameter estimates of the BRD are obtained using the method of maximum likelihood estimation.

The Weibull-Inverse Weibull Composite Distribution for Modeling Reliability Data

✱Kahadawala Cooray, University of Nevada, Las Vegas, Department of Mathematical Sciences, Las Vegas, NV 89154, *coorayk@unlv.nevada.edu*; Sumith Gunasekera, University of Nevada, Las Vegas; Malwane Ananda, University of Nevada, Las Vegas

Key Words: Reliability modeling, Weibull density, Inverse Weibull, Composite Distributions

The Weibull distribution, which is frequently used for life data analysis, is composited with inverse Weibull distribution to obtain a flexible, reliable, long-tail parametric distribution for modeling reliability data. This smooth and continuous natural composition of Weibull and inverse Weibull family is a two-parameter Weibull density up to an unknown threshold value, and a two-parameter inverse Weibull density for the remainder. The resulting two-parameter density is similar, in shape, to the Weibull density, yet its upper tail is larger than the Weibull density, and the tail behavior is quite similar to the inverse Weibull density. The least square parameter estimation techniques of this composite family are discussed by providing well-known examples.

Inference Models for Time-Varying, Stress-Accelerated Life Tests Using Generalized Linear Models

✱Jinsuk Lee, Arizona State University, Dept. of Industrial Engineering,, Arizona State University, Tempe, AZ 85287, *jinsuk.lee@asu.edu*; Rong Pan, Arizona State University

Key Words: time-varying stress accelerated life test, censoring, generalized linear models

In this paper we present a statistical inference model for a time-varying stress accelerated life test using GLMs. We provide a general time-varying stress ALT model for exponential failure data. Through the proportional hazard model the likelihood function of complicated time-varying stress ALT is transformed to a Poisson distribution with censoring variate and the life-stress relationship is defined by a log link function of GLM. For the inference procedure the parameters of the log-linear model are estimated via the MLE method. IRWLS is used to implement the Newton-Raphson method for iterative solutions and obtain the MLEs. In addition, Bayesian analysis using Jeffreys' prior is performed via the MCMC simulation technique. A case study is presented for the validation and illustration of the model. The model using GLMs is suitable for time-varying stress ALT and easy to implement.

Active Component Identification in Reliability Modeling for Complex Systems

✱Scott Vander Wiel, Los Alamos National Laboratory, MS F600, Los Alamos, NM 87545, *svanderwiel@comcast.net*; Earl Lawrence, Los Alamos National Laboratory

Key Words: reliability, prior, system, component, variable selection

Standard priors for component reliabilities can lead to overly pessimistic inference at the system level when pass/fail test data are sparse. In part, this problem is caused by poor separation between inactive components that operate nearly flawlessly and active components that negatively impact the system reliability. We explore the use of a hierarchical mixture of negative log-gamma (NLG) distributions as priors for components in a series system. This has two benefits. First, the properties of the NLG ensure an appropriate induced prior at the system level. Second, the hierarchical prior (inspired by Bayesian variable selection) more effectively separates active and inactive components. Extensions to modeling reliability over time are also discussed.

482 Bayesian Methods

Section on Bayesian Statistical Science, Section on Statistical Computing

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Bayesian Evaluation of Surrogate Endpoints Using Power Priors

✱Chunyao (Amy) Feng, Amgen, Inc, One Amgen Center Dr., Thousand Oaks, CA 91362, *afeng@amgen.com*; John W. Seaman, Jr., Baylor University; Stacy Lindborg, Eli Lilly and Company

Key Words: Bayesian Analysis, Surrogate Endpoint, Power Prior

To save time and reduce the size and cost of clinical trials, surrogate endpoints are frequently measured instead of true endpoints. The proportion of the treatment effect explained by surrogate endpoints (PTE) is a widely used, albeit controversial, validation criteria. As an alternative to the use of PTE, we develop a Bayesian model for relative effect and the association between surrogate and true endpoints. Three prior structures are used: diffuse normal, informative, and power priors. The latter two prior structures are based on historical data. The power priors, developed by Ibrahim and Chen (2000) are used to attenuate the influence of the historical data on the current likelihood.

Bayesian Case Influence Measures and Applications

✱Hyunsoon Cho, The University of North Carolina at Chapel Hill, 3920 South Roxboro Street, Apt213, Durham, NC 27713, *hscho@bios.unc.edu*; Hongtu Zhu, The University of North Carolina at Chapel Hill; Joseph G. Ibrahim, The University of North Carolina at Chapel Hill

Key Words: Case influence measures, Markov chain Monte Carlo, Model complexity, Model selection criterion, Bayesian regression models

We introduce three types of Bayesian case influence measures based on case deletion, and propose Bayesian information criterion and goodness of fit statistics. We examine the asymptotic approximations and equivalencies of the proposed measures. We show that the sum of the proposed Bayesian case-deletion measures can measure model complexity and are associated with the effective number of parameters in deviance information criterion. We construct a Bayesian information criterion using the posterior mean of the expected log likelihood. We show that the proposed measure of model complexity can asymptotically correct the asymptotic bias of the posterior mean of the log likelihood. The proposed goodness-of-fit statistics is based on the sum of the Bayesian case-deletion measures. We illustrate the methodology using theoretical as well as numerical examples in Bayesian regression models.

From Semi to Fully Bayes Factors in Hypothesis Testing

✱Kai Fun Yu, National Institutes of Health, Bldg 6100, Room 7B05, National Institutes of Health, Bethesda, MD 20892-7510, *yukf@mail.nih.gov*; Albert Vexler, National Institute of Child Health and Human Development; Chengqing Wu, Yale University

Key Words: Likelihood ratio, maximum likelihood, Bayes factor, hypothesis testing, signifcance level, most powerful

We introduce a principle for statistical hypothesis testing which is a hybrid of the two classical approaches of likelihood ratio test and Bayes factor. We examine the properties of the tests derived from this principle, including some optimality properties. We demonstrate that the new testing method can be executed and applied to practical studies. Under some conditions, the classical requirement that the significance level of the proposed test is controlled. This new method can be a compromise between the two classical approaches.

Effect of Hyperparameters in the Normal Conjugate Model on Posterior Estimates

✱Susan Alber, The University of Texas M.D. Anderson Cancer Center, Dept. of Biostatistics Unit 447, 1515 Holcombe Blvd., Houston, TX 77030, *Bayesiology@yahoo.com*; J. Jack Lee, The University of Texas M.D. Anderson Cancer Center

Key Words: prior specification, Bayesian, sensitivity analysis

In Bayesian analysis prior specification is critical. We investigate the frequentist properties of point estimates in a normal conjugate model for multiple treatment groups with correlation between treatment means. Assuming that the true distribution of the data is normal with independent observations, given fixed expected value for each treatment, and a common variance across treatments, we demonstrate how the mean square error (MSE) of point estimates for treatment means and variances depend on the choice of the hyperparameters. We use our results to provide guidelines for choosing hyperparameters in data analysis, where the true data generating model is not known. We conclude that for the normal model it is feasible to choose conservative hyperparameter values that reduce the MSE of parameter estimates as compared to the standard frequentist estimates.

Bayesian Inference for the Squared Multiple Correlation Coefficient

✱Joseph F. Lucke, The University of Texas Health Science Center at Houston, Center for Clinical Research and Evidence-based Medicine, 6431 Fannin, MSB2.106, Houston, TX 77030, *Joseph.F.Lucke@uth.tmc.edu*

Key Words: Bayesian inference, squared multiple correlation, posterior expectation, hypergeometric functions

The sample squared multiple correlation coefficient is well-known to be an upwardly biased estimator of its population parameter, especially when the sample size is small and the number of predictors approaches the sample size. A small cottage industry has undertaken to correct its bias and has produced a number of adjusted estimators, the most popular being Fisher's. Curiosity prompts the question whether a Bayesian estimator, namely the posterior expectation, suffers from an analogous defect. Finding out involves the hypergeometric and the generalized hypergeometric functions. As it turns out, the Bayesian estimator is automatically 'adjusted' for small samples with many predictors. Performance comparisons between the frequentist and Bayesian estimators are presented.

Unifying the Named Natural Exponential Families and Their Relatives

✱Kari F. Lock, Harvard University, 102 Line St. #1, Somerville, MA 02143, *klock@fas.harvard.edu*; Carl Morris, Harvard University

Key Words: Normal, Poisson, Gamma, Binomial, Pearson families, Quadratic variance function

Five of the six univariate natural exponential families (NEF) with quadratic variance functions (QVF), meaning their variances are at most quadratic functions of their means, are the Normal, Poisson, Gamma, Binomial, and Negative Binomial distributions. The sixth is the NEF-CHS, i.e. the NEF generated by convolved Hyperbolic Secant distributions. This talk unifies these distributions and their relatives, including Pearson's family of conjugate distributions (Inverted Gamma, Beta, F, and Skew-t), conjugate mixtures (including the Polya I and Polya II binary urn schemes), and conditional distributions (including the Hypergeometric and Negative Hypergeometric).

483 Topics in Statistical Estimation

IMS

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Admissibility of Generalized Bayes Estimators Through Markov Chain Arguments

✱Brian Shea, The University of Minnesota, 313 Ford Hall, 224 Church Street SE, Minneapolis, MN 55455, *bshea@stat.umn.edu*

Key Words: Admissibility, Formal Bayes, Markov chain, Recurrence, Multivariate normal distribution

Given a parametric model and improper prior distribution, Eaton (1992 {\it Annals}, 1999 {\it PNA}) provided conditions under which recurrence of a Markov chain is a sufficient condition for admissibility of the generalized Bayes estimator under squared error loss. Eaton {\it et al} (2007, {\it Annals} to appear) provide a method of reducing the Markov chain to one dimension as well as moment conditions for the reduced chain's transition kernel that guarantee admissibility. Their results apply to estimating a bounded function of the parameter. We extend these results to the case of estimating unbounded functions of the parameter, and the important special case of estimating the mean of a p-dimensional multivariate normal distribution is considered. Generalized Bayes estimators of the mean arising from a class of improper priors are shown to be admissible under squared error loss.

Doubly Smoothed Maximum Likelihood Estimation

✱Byungtae Seo, Texas Tech University, Department of Mathematics and Statistics, Box 41042, Lubbock, TX 79409-1042, *byungtae.seo@ttu.edu*; Bruce G. Lindsay, The Pennsylvania State University

Key Words: Consistency, Minimum distance, kernel smoothing, Maximum likelihood

In some models, both parametric and not, maximum likelihood estimation fails to be consistent. We show this failure of ML method with some examples and notice the paradox that, in those same models, maximum likelihood estimation would have been consistent if the data had been measured with error. With this motivation we define doubly smoothed maximum likelihood as a natural mechanism for adding measurement error. We show the proposed estimation procedure gives universal consistency in independent and identically distributed data. Some practical guidelines for the choice of kernel and tuning parameter are given. A Monte Carlo computational method is also discussed.

The Perfect Median Revisited

✱Winston Richards, The Pennsylvania State University, *ugu@psu.edu*

Key Words: Perfect Median, Pell Numbers

In the Fall 2005 issue of the Emissary, News Letter of the Mathematical Sciences Institute, David Gale posed this interesting problem and I set about attempting a solution. The Problem: Say that an ordered sequence of real numbers has a perfect median if there is a term such that the sum of the terms preceding it is equal to the sum of the terms following it. For instance, 6 is a perfect median of the sequence 1, 2, 3, 4, 5, 6, 7, 8. For which n does the counting sequence 1, 2,., n have a perfect median? My approach that led to the discovery of a new formula in terms of the Pell Numbers for determining the Perfect Median will be presented.

Folded Parametric Families

✱Malwane Ananda, University of Nevada, Las Vegas, Department of Mathematical Sciences, Las Vegas, NV 89154, *malwane.ananda@unlv.edu*; Kahadawala Cooray, University of Nevada, Las Vegas; Sumith Gunasekera, University of Nevada, Las Vegas

Key Words: Folded Distributions, Cauchy Distribution, Laplace Distribution

In some practical applications, measurements are recorded without their algebraic sign. As a consequence, the underlying distributions of measurements are replaced by distributions of absolute measurements, and the resulting distributions are known as folded distributions. In general, folded distributions are positively skewed and have non-zero density value. Therefore, these distributions are useful to analyze the data sets with zero data points. The folded normal and folded logistic distributions and their applications have already been discussed in detail in statistical literature. This paper is confined to discuss some properties of the folded Cauchy and folded Laplace distributions. Parameter estimation techniques are discussed and the advantages of using these distributions are demonstrated using well-known examples.

Least Squares Estimation of Density of Very Large Prime Numbers

✱Arthur Cohen, Nassau Community College, 55 sandy hollow rd., Port Washington, NY 11050, *art1899@aol.com*

Key Words: least squares estimate, density of prime numbers, very large numbers, evaluation of extrapolation, Fundamental prime no. theorem, Gauss' Logarithmic Integral of N

Based on the actual number of prime numbers up to 10 to the 15th power, a least squares estimate is developed which relates the density of primes to the size of its associated order of magnitude interval. This estimate is exceedingly good. Since the actual data ends with a density of about 3%, I extrapolated to find that interval where the density is approximately 1%, and then .1%. The former came to a 43 digit number, and the latter came to a 390 digit number. These estimates were then compared to estimates generated using the Fundamental Prime Number Theorem as well as the first two terms of the series based on Gauss' Logarithmic Integral of N. These latter estimates agreed closely with my 1% estimate, but differed by about 10% with my .1% estimate; that is, they obtained a 435 digit number. Finally, for extremely large primes N, we show that the density approaches (1/ln N).

Convergence Rates for Variable-at-a-Time Metropolis-Hastings Algorithms

✱Ronald Neath, Baruch College, City University of New York, Department of Statistics and CIS, One Bernard Baruch Way, Box B11-220, New York, NY 10010, *Ronald_Neath@baruch.cuny.edu*; Galin Jones, The University of Minnesota

Key Words: Convergence rate, Geometric ergodicity, Markov chain Monte Carlo, Metropolis-Hastings algorithm, Minorization condition, Regenerative simulation

In a Metropolis-Hastings independence sampler, proposals are drawn from a candidate distribution that does not depend on the current state of the chain. Identifying a suitable candidate density can be a challenging proposition in high-dimensional problems. A popular modification is to update the chain one component at a time, effectively replacing a p-dimensional problem with p one-dimensional problems. Of course the resulting algorithm is no longer an independence sampler, and thus results on uniform ergodicity (Mengersen and Tweedie, 1996) and regeneration times (Mykland, Tierney and Yu, 1995) do not apply. Two main problems are considered in this talk: (i) establishing a bound on the rate of convergence to the stationary distribution, and (ii) the use of regenerative simulation as a means of estimating Monte Carlo standard errors.

484 Time Series Modeling

Business and Economics Statistics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Simple ARFIMA Approximation to the Limiting Aggregate Structure of Long Memory Process

✱Kasing Man, Western Illinois University, Dept. of Information Systems and Decision Sciences, College of Business and Technology, Macomb, IL 61455, *KS-Man@wiu.edu*

Key Words: Forecasting, Long memory process, Temporal aggregation, Time Series

This article studies the precision of Man and Tiao's (2006) low-order ARFIMA (0,d,d*) approximation to Tsai and Chan's (2005) limiting aggregate structure of long memory process. The latter can be regarded as an extension of fractional Gaussian noise. In computing autocorrelations over long lags using the exact formula for larger value of d, it is noted that numerical problem might occur. The use of (0,d,d*) model thus provides a feasible alternative as a really close approximation. We also derive a formula to approximate the Gamma function as a means to assess the closeness of the approximation. In forecasting future aggregate, we demonstrate the close performance of using the (0,d,d*) model and the exact aggregate structure for positive d. In practice, this provides a justification of using low-order ARFIMA model in predicting future aggregates of long memory process.

Computationally Efficient Gaussian Maximum Likelihood Methods for Vector ARFIMA Models

✱Rebecca J. Sela, Stern School of Business/New York University, 44 West Fourth Street, New York, NY 10012, *rsela@stern.nyu.edu*; Clifford Hurvich, Stern School of Business/New York University

Key Words: long memory, multivariate time series, maximum likelihood, computational methods, ARFIMA, VARFIMA

In this paper, we discuss two distinct multivariate time series models that extend the univariate ARFIMA model. We describe algorithms for computing the covariances of each model, for computing the quadratic form and approximating the determinant for maximum likelihood estimation, and for simulating from each model. We compare the speed and accuracy of each algorithm to existing methods and measure the performance of the maximum likelihood estimator compared to existing methods. We also fit models to data on unemployment and inflation in the United States and to data about precipitation in the Great Lakes.

Nonlinear Exponential Smoothing and Positive Data

✱J. Keith Ord, Georgetown University, 37th and "O" Streets NW, Washington, DC 20057, *ordk@georgetown.edu*; Muhammad Akram, Monash University; Rob J. Hyndman, Monash University

Key Words: forecasting, time-series, exponential smoothing, positive valued processes, state space models

We consider nonlinear exponential smoothing state space models that are used to describe non-negative observations. When the innovations process is assumed to be Gaussian, the resulting prediction distribution may have an infinite variance beyond a certain forecasting horizon. Further, such processes may converge almost surely to zero. We then explore effects of using an (invalid) Gaussian distribution to describe the innovations process when the underlying distribution is lognormal. Our results suggest that this approximation causes no serious problems for parameter estimation or for forecasting one or two steps ahead. However, for longer-term forecasts the true prediction intervals become increasingly skewed. The performance of the Gaussian approximation is compared with those of two lognormal models for short-term forecasting using data on weekly jewelry sales.

Integer-Valued Time Series and Renewal Processes

✱Yunwei Cui, Clemson University, 29634, *ycui@clemson.edu*; Robert Lund, Clemson University

Key Words: Integer-valued Time Series, Renewal Processes, superpositioning stationary renewal processes, AR(1)

This talk proposes a new model for time series of integer counts. The model is built by superpositioning stationary renewal processes. The model produces stationary series whose marginal distribution at any fixed time can be binomial. We first describe the model and its statistical properties.

Attention then turns to relating the model to autoregressive moving-average time series, the AR(1) model with positive autoregressive coefficient being key. Examples and simulation results are given.

A Modified Approach to Obtaining Sieve Bootstrap Prediction Intervals for Time Series

✱Purna Mukhopadhyay, Missouri University of Science and Technology, Department of Mathematics and Statistics, 400 W 12th Street, 202 Rolla Building, Rolla, MO 65409-0020, *pm442@mst.edu*; V. A. Samaranayake, Missouri University of Science and Technology

Key Words: ARMA processes, Forecast intervals, Residual resampling, Invertible processes

The Sieve Bootstrap approach to obtaining prediction intervals for invertible ARMA processes employs a finite autoregressive approximation to the time series to obtain residuals for re-sampling and to generate a bootstrap distribution of future values. The advantage of this approach is that it does not require the knowledge of the orders p and q associated with the ARMA process. The coverage probabilities of prediction intervals based on this method, however, are liberal and in some cases falls substantially below nominal levels. We propose several modifications to the above method in order to obtain better coverage. These modifications aim to capture the sampling variations that are ignored by current methods. Monte Carlo studies show that the proposed method outperforms the standard Sieve bootstrap as well as other procedures in most cases.

Forecasting in Linear Autoregressive Models with Heteroscedastic Measurement Error

✱Yorghos Tripodis, University of Massachusetts, 710 North Pleasant Street, Department of Mathematics & Statistics, Amherst, MA 01003, *tripodis@math.umass.edu*

Key Words: measurement error, heteroscedasticity, Kalman Filter

Conditional heteroscedasticity has often been used in modeling and understanding statistical data. A variable of interest is observed with some measurement error with non-constant variance. The measurement error variance is a known function of the unobserved variable. This is a common case, especially when the sampling effort is a function of the non-constant population level. We propose a two step estimation method. The first step consists of fitting a constant variance model and uses a residual-based estimator for the heteroscedastic measurement error variance. A correction is applied to adjust for the erroneous use of constant variance. We then estimate the process variance using maximum likelihood. We look at the behavior of these estimators for small and large samples. Applications with real data illustrate the usefulness of the proposed techniques.

A Reformulation of Generalized Least Squares Estimators in Autocorrelated Regression

✱Jaechoul Lee, Boise State University, 1910 University Dr, Department of Mathematics, Boise, ID 83725-1555, *jaechlee@math.boisestate.edu*

Key Words: Autocorrelation, Efficient computing, Generalized least squares, Ordinary least squares, Simple regression

This paper studies generalized least squares estimators in simple linear regression with serially correlated errors. Closed-form expressions for the generalized least squares estimation are presented for a given inverse covariance matrix with general stationary structure. With the linear trend plus autoregressive moving-average error regression model, the presented formulations produce further explicit expressions of the generalized least squares trend estimator and variance. As an application of these reformulated expressions, a new generalized least squares computation method that reduces the effort of inverting the covariance matrix but produces an equivalent estimate and variance is developed. A new sufficient and necessary condition that the generalized least squares estimators are the same as the ordinary least squares estimators is also provided in a closed form.

485 Survival Analysis II: New Developments

Biometrics Section, Section on Nonparametric Statistics, Biopharmaceutical Section, Section on Quality and Productivity

Thursday, August 7, 8:30 a.m.–10:20 a.m.

A Parametric Method for Testing Proportional Odds Assumption

✱Qing Xu, U.S. Food and Drug Administration, 10903 New Hampshire Avenue,Bldg. #22, Room1213, DBV, CDER, Silver Spirng,, MD 20993-0002, *qing.xu@fda.hhs.gov*; Jong-Hyeon Jeong, University of Pittsburgh

Key Words: proportional odds, survival data, parametric method

We develop a parametric method for detecting the validity of the proportional odds assumption between two groups for survival data. This study was based on the premise that the test procedures taking advantage of knowledge of the distributional information about the data will improve the sensitivity of an existing nonparametric test method. In the proposed method, a parametric form of the weight function is supplemented into an existing nonparametric test statistic for testing the proportional odds assumption. We evaluate type I error probabilities and power of the new method by using the simulated survival data following the log-logistic distribution. The results indicate that the parametric test performs with a higher sensitivity than the existing one, as expected, providing an alternative method of testing the proportional odds assumption often encountered in clinical trial data.

Semiparametric Cure Rate Models with Random Effects

✱Guoqing Diao, George Mason University, Department of Statistics, MSN 4A7, 4400 University Dr., Fairfax, VA 22030, *gdiao@gmu.edu*

Key Words: Box-Cox transformation, cure fraction, mixture cure model, proportional hazards cure model

We propose a novel class of cure rate models for multivariate failure time data with a survival fraction. The class is formulated through a transformation on the unknown population survival function. It incorporates random effects to account for the underlying correlation, and includes the mixture cure model structure and the proportional hazards cure model structure as two special cases. We show that the nonparametric maximum likelihood estimators (NPMLE) for the parameters of these models are consistent and asymptotically normal. The limiting variances achieve the semiparametric efficiency bounds and can be consistently estimated. Simulation studies demonstrate that the proposed methods perform well in practical situations. This class of models is illustrated with a real example.

Estimating the Convolution of Distributions Under the Partial Koziol-Green Model of Random Censorship

✱Ke Wu, California State University, Fresno, 5245 North Backer Avenue M/S PB 108, Dept. of Maths., California State Univ., Fresno, Fresno, CA 93740-8001, *kewu@csufresno.edu*

Key Words: convolution, partial Koziol-Green model, random censorship, influence function, partial ACL estimator, product limit estimator

In survival analysis, a model of informative censoring is the Koziol-Green (KG) model, where the survival function of the censoring times is some power of the survival function of the lifetimes. A generalization of the KG model is the partial Koziol-Green (PKG) model which allows the lifetimes to be censored by two types of variables, one of which censors in an informative way and the other one in a non-informative way. In this paper, an estimator is proposed for the convolution of distribution functions under the PKG model. Our estimator uses the partial ACL estimator for the survival function under the PKG model instead of the product limit estimator of Kaplan and Meier (1958) that is used in Lagakos and Reid (1981). The asymptotic distribution of the new estimator is established.

A Finite Mixture Survival Model To Characterize Risk Groups of Neuroblastoma

✱Sally Hunsberger, National Institutes of Health, 6130 Executive Blvd, rm 8120, Rockville, MD 20852, *sallyh@ctep.nci.nih.gov*; Paul S. Albert, National Institutes of Health; Wendy B. London, University of Florida

Key Words: parametric cure model, classification

Neuroblastoma is a childhood cancer with patients experiencing heterogeneous survival outcomes despite aggressive treatment. Due to this heterogeneity, it is of interest to identify patients with similar types of neuroblastoma. Oncologists are especially interested in identifying patients who will be cured so that the minimum amount of a potentially toxic treatment can be given to this group of patients. We analyze a large cohort of neuroblastoma patients and develop a cure mixture model that uses covariates to predict the probability of being in a cure group or other (one or more) risk groups. A prediction method is developed that uses the estimated probabilities to assign a patient to different risk groups. The robustness of the model and the prediction method is examined via simulation by looking at misclassification rates under misspecified models.

Selecting Representative Trees in Random Forest for Survival Data

Mousumi Banerjee, The University of Michigan; Ying Ding, The University of Michigan; ✱Anne-Michelle Noone, Georgetown University

Key Words: tree-based methods, survival data, random forest, out-of-bag error, similarity metric

Tree-based methods are popular tools for prognostic stratification. Ensemble techniques such as random forest improve accuracy in prediction and address instability in a single tree. However, individual trees are lost in the forest. In this paper, we propose a methodology for selecting the most representative trees in a forest for survival data, based on three tree similarity metrics. For any two trees, the metrics are chosen to measure similarity of the covariates used to split the trees; reflect similar clustering of patients in the terminal nodes; and measure similarity in predictions. The most representative trees in the forest are chosen based on the average similarity score assigned to each tree. Out of bag estimates of error are computed for the most representative trees using a neighborhood of similar trees. Finally we illustrate the methods using a breast cancer data set.

Partially Functional Quantile Regression Model for Survival Data

✱Jing Qian, Emory University, 1518 Clifton RD NE 3rd FL, Atlanta, GA 30322, *jqian@emory.edu*; Limin Peng, Emory University

Key Words: Censoring, Empirical Process, Martingale, Regression Quantiles, Resampling

Quantile regression has received increased attention in survival analysis because of its good interpretability and great flexibility. In recent work by Peng and Huang (2008), a new censored quantile regression approach has been developed by utilizing the inherent martingale structure of survival data, without requiring stringent censoring assumption and involving complicated algorithms. In this paper we further extend Peng and Huang (2008)'s technique to partially functional quantile regression model, which in reality is expected to achieve a better balance between efficiency and robustness than the standard fully model. We establish the asymptotic properties of the resultant estimators and develop a simple resampling inference procedure. The finite-sample performance of the proposed method is evaluated via simulation studies and illustrated with an application to a dialysis study.

486 Statistical Methods for Genome-Wide Association Studies II

Biometrics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Statistical Methods for Detecting Deletions Using SNP Genotypes in Case-Control Studies

✱Chih-Chieh Wu, M.D. Anderson Cancer Center, Houston, TX 77030, *ccwu@mdanderson.org*; Sanjay Shete, The University of Texas M.D. Anderson Cancer Center; Bo Peng, M.D. Anderson Cancer Center; Jianzhong Ma, M.D. Anderson Cancer Center; Christopher I. Amos, M.D. Anderson Cancer Center

Key Words: genetic, case-control studies, deletion, genotype

Recent studies show that genomic deletions can play an important and crucial role in genetic basis of complex traits. Existing methods for identifying genomic deletions from genetic markers are generally based on the observation of Mendelian transmission failure in parents-child trios, Hardy-Weinberg disequilibrium, and null genotypes. We propose statistical methods that use SNP genotype data to detect genomic deletions and to assess the association with the disease of interest. One approach assesses the excess of homozygosity of contiguous genetic markers in case subjects compared with those in control subjects. The second approach tests for the run of homozygosity in case-series studies. SNP marker data with and without linkage disequilibrium were simulated from chromosome 2q (80cM–180cM) of the HapMap. Analyses for simulation studies will be presented.

Family-Based Conditional Gene Test

✱Thomas Hoffmann, Harvard School of Public Health, 655 Huntington Ave., Department of Biostatistics, Boston, MA 02115, *tjhoffm@fas.harvard.edu*; Nan Laird, Harvard School of Public Health

Key Words: family-based, genetics, association test

Consider the situation in which we have a candidate region of interest, in which the markers are in LD, and testing for association of the main effects of some of the markers are significant. We would like to be able to test if one of these markers is the DSL. The effect size is a function of allele frequency and LD, this it is often not enough to simply find the maximum univariate chi-squared test. We introduce a family-based multi-marker conditional gene test to address this question for binary or continuous traits in nuclear families.

A Bayesian Change-Point Algorithm for the Analysis of SNP Data

✱Fridtjof Thomas, University of Tennessee Health Science Center, Memphis, TN 38105, *fthomas4@utmem.edu*; Stanley B. Pounds, St. Jude Children's Research Hospital

Key Words: microarrays, SNP, genomics, copy number alteration, change-point, Bayesian

High-resolution genomics data in the form of single nucleotide polymorphism (SNP) arrays can be used in a paired data context to compare cancer tissue to normal samples in an effort to identify regions of genomic amplification or deletion. Such regions potentially contain oncogenes or tumor suppressor genes and are therefore of particular interest. We apply here a Bayesian change-point algorithm to pre-normalized signals from SNP microarrays obtained from a set of leukemia samples in an effort to infer regions of copy number alteration. This algorithm detects multiple change-points where a change can be in the mean of the subsequent measurements, in their variance, in their autocorrelation structure, or in a combination of two or all of these aspects.

Imputing Missing Data in Case-Parent Triad Studies

✱Tracy L. Bergemann, The University of Minnesota, 420 Delaware St SE, MMC 303, Minneapolis, MN 55455, *berge319@umn.edu*

Key Words: Case-parent triads, ZIP regression, Genetic epidemiology, Missing data

Case-parent triad studies sample affected offspring, manifesting a disease or phenotype of interest, as well as their parents. If parents of the offspring are alive and consent to genotyping, the use of nuclear families can be a powerful way to measure association. For some triads, not all genotypes are available for each family member, usually the paternal DNA, at the time of the study. The EM algorithm is widely used to impute missing genotypes. Some of these algorithms under-estimate the variance of the regression coefficients, especially when the minor allele frequency is less than 0.10 or when the degree of missingness is large (> 20%). I propose a new method by solving a Zero-Inflated Poisson (ZIP) likelihood directly. This likelihood simultaneously estimates allele frequencies, regression coefficients and imputes missing data to reduce bias and type I error, and maintain power.

Statistical Methods for Inferring Duplicates Among Genotyped Samples

✱Lei Shen, Eli Lilly and Company, Lilly Corporate Center, DC 0734, Indianapolis, IN 46285, *shen_lei@lilly.com*

Key Words: SNP, allele frequency, duplication, sample handling error, genotyping error, correlation

In an analysis involving genotype information from a large number of SNPs, it is often important to check for duplicate samples, especially if data are combined across studies. Depending on the cause of duplication, this information can improve the validity of downstream analyses by eliminating sample handling errors or accounting for intra-subject correlations when necessary. Furthermore, combining and cross-checking duplicate samples can improve data quality and provide information on assay performance. We consider several issues inherent to the inference of duplicate samples such as missing data, genotyping errors, correlations between SNPs, and incorporation of allele frequency information. A number of methods are proposed and compared in terms of applicability, robustness, and ability to correctly identify duplicate samples from various causes.

Rapid Genotype Imputation and Analysis of Resequencing Data Using Markov Models

✱Yun Li, The University of Michigan, 1071 Barton Dr Apt205, Ann Arbor, MI 48105, *ylwtx@umich.edu*; Cristen Willer, The University of Michigan; Jun Ding, The University of Michigan; Paul Scheet, The University of Michigan; Goncalo Abecasis, The University of Michigan

Key Words: genotype imputation, GWAS, Markov model, meta-analysis, resequencing, gene mapping

I propose and implement efficient Markov models for genome wide association studies (GWAS), a more powerful tool to detect genetic variants with small individual contributions to complex traits. Specifically, I consider how to impute several million common SNPs not typed in a GWAS. Imputation-enabled meta-analyses identified multiple novel loci influencing risk of diabetes, coronary artery disease, height, or lipid levels. I also consider how to obtain accurate estimates of individual sequences from shotgun resequencing data. My method allows the more cost-effective more-individual-lower-coverage design both for generating a public resequencing database and for capturing variants in individual resequencing based GWAS. My approach should stimulate the advent of large-scale resequencing based GWAS and foster the detection of rare variants not adequately assessed with current approaches.

An Empirical Bayesian Method To Correct for Winner's Curse in Genetic Association Studies

✱Rui Xiao, The University of Michigan, Department of Biostatistics, 1420 Washington Heights, Ann Arbor, MI 48105, *xiaor@umich.edu*; Michael Boehnke, The University of Michigan

Key Words: genomewide association studies, genetic effect estimate, winner's curse, empirical bayesian

Genetic association mapping is a powerful method to detect genetic variants that predispose to human disease. Investigators are also interested in estimating the genetic effect on disease risk of each identified variant. Initial positive findings of the genetic effect estimate tend to be upwardly biased particularly if they were the first to reach the statistical significance level, a phenomenon known as the winner's curse. Overestimation of genetic effect size in initial studies may cause follow-up studies to be underpowered and so to fail. In this paper, we propose an empirical Bayesian method to correct for the overestimation. Our method incorporates information from genomewide association studies as a prior distribution for the genetic effect size throughout the genome, and then combines the locus specific data to reduce the bias.

487 Statistical Issues with Genetic Data ●▲

Biopharmaceutical Section, Biometrics Section
Thursday, August 7, 8:30 a.m.–10:20 a.m.

Statistical Methods for the Analysis of RNAi Screens

✱Imola K. Fodor, Genentech, Inc., 1 DNA Way, MS 44-1B, South San Francisco, CA 94080-4990, *fodor.imola@gene.com*; William F. Forrest, Genentech, Inc.

Key Words: RNAi, siRNA, high throughput screening, pathway, normalization, mixed-effects model

By selectively turning off one gene at a time, cell-based RNA interference (RNAi) screens provide an efficient means to investigate the genes involved in biological pathways of interest. This relatively new technology is currently being used both to conduct basic research on intracellular signaling pathways and to discover potential new drug targets. RNAi experiments typically consist of data from many microtiter plates, and statistical analyses can span from classical statistical tools (such as data transformations and mixed effects modeling) through areas of active statistical research (such as multiple testing frameworks). We describe some of the statistical challenges in RNAi experiments and illustrate some solutions through a case study.

Better Prognosis by Multiple-Gene Classifier with Pair Information

✱Shuyan (Sabrina) Wan, Merck Research Laboratories, RY34-A316, 126 East Lincoln Avenue, Rahway, NJ 07065, *shuyan_wan@merck.com*; Xiang Yu, Merck Reserach Laboratories; Peggy Wong, Merck Research Laboratories

Key Words: Random Forest, biomarker, disease prognosis, pair information,

Molecular profiling techniques for disease prognosis have been actively applied in biomedical research recently. When clinical information is available, it is attractive for medical researchers to match patients from two disease (or treatment) groups into pairs to increase the power to detect differentially expressed genes. For such design, it is necessary to take into account of the correlation between genes and between samples to build a multiple-gene classifier. Thus we proposed a Random Forest (RF) method conditional on matching clinical covariates to select gene signature. Simulation studies showed that the RF method was able to build a composite biomarker classifier and predict the incoming patient's group accurately. The method was also applied to a blinded dataset in cardiovascular disease and the results were promising.

Domain-Enhanced Analysis Using the Gene Ontology with Different Classification Response Variables

✱Jiajun Liu, Merck Research Laboratories, RY34-A316, P.O. Box 2000, 126 East lincoln Avenue, Rahway, NJ, 07065, NJ 07065, *jiajun_liu@merck.com*; Jacqueline Hughes-Oliver, North Carolina State University; Jason Osborne, North Carolina State University; Jack A. Menius, GlaxoSmithKline

Key Words: Gene Expressions, Gene Aggregation, Gene Ontology, Gene Set Enrichment, Partial Least Square

New biological systems technologies give scientists abilities to measure thousands of biomolecules including genes, proteins, lipids, and metabolites. We use a procedure called Domain Enhanced Analysis (DEA), a "top-down" approach to perform domain aggregation by first combining gene expressions using the Gene ontology before testing for differentially expressed patterns; the benefits are greater sensitivity for detecting signals. In original DEA procedure, the first scores from the Partial Least Square procedure are used to test for differentially expressed patterns using the t test. We proposed new tests by finding the true null distribution of each PLS score adjusted for the size of the GO term to replace the original t test. We also discuss the impact of our testing procedure with different coding of our classification response variable, namely 0/1 or -1/1 for data with two classes.

Gene-Class Testing for Multigroup Comparisons

✱Ching-Wei Chang, National Center for Toxicological Research, 72079, *ching-wei.chang@fda.hhs.gov*; James J. Chen, National Center for Toxicological Research

Key Words: Gene expression, generalized linear model, random effect

Gene-class testing (GCT) has been proposed for gene expression analysis to understand biological functions. In GCT, the hypothesis is that if there are treatment effects in the gene class. Some statistical methods for gene-class testing are only for one-sided test, that is, the changes of individual gene expressions in a gene class are all in one direction: either up or down. The one-side test might not close to the reality. Another problem is that current methods only focus on the two-group comparison. However, many studies involve comparisons of more than two groups. In this study, we use the generalized linear model by assuming random effects of gene expressions for a two-sided test in multi-group comparisons. The generalized linear model approach can take the clinical covariates into consideration. We propose computing techniques to deal with the large number of parameters.

Statistical Test on Nonrandom Clustering with Application to Genetic Somatic Mutation

✱Jingjing Ye, Pfizer, Inc., 10777 Science Center Drive, San Diego, CA 92121, *Jingjing.Ye@pfizer.com*; Elizabeth A. Lunney, Pfizer, Inc.; Paul A. Rejto, Pfizer, Inc.; Adam Pavlicek, Pfizer, Inc.; Chi-Hse Teng, Pfizer, Inc.

Key Words: nonrandom clustering, order statistics, somatic mutation, oncogenes, tumor suppressor, COSMIC (Catalog Of Somatic Mutations In Cancer)

Genetic somatic mutations are base-pair alterations in DNA in any cells of the body except germ cells. Previous studies have shown somatic mutations are enriched in cancer cells, thus detection of the mutations is crucial for targeted therapies in cancer care. Nonrandom occurrence of mutation, without presence of obvious hotspots is strongly indicative of potential selection for cancer. A new and rapid statistics test for detection of nonrandom mutation clustering is derived by forming mutation positions as order statistics. Our test statistics are the differences between pair-wise order statistics, which indicates mutation cluster size. The p-values are derived from the theoretical and approximate distributions of the test statistics. Using data in the COSMIC database, we have validated our method by detecting well-known clusters in oncogenes and in tumor suppressors.

488 Modeling Survival Data ●▲

Biopharmaceutical Section, Section on Nonparametric Statistics, Section on Quality and Productivity, Biometrics Section

Thursday, August 7, 8:30 a.m.–10:20 a.m.

Covariate-Adjusted Nonparametric Survival Curve Estimation

✱Honghua Jiang, Eli Lilly and Company, *jianghh@lilly.com*; Jame Symanowski, Nevada Cancer Institute; Yongming Qu, Eli Lilly and Company; Yanping Wang, Eli Lilly and Company

Key Words: Survival curve, non-parametric, adjusting for covariates

Kaplan-Meier survival curve estimation is a commonly used non-parametric method to evaluate survival distributions for groups of patients in the clinical trial setting. However, this method does not permit covariate adjustment which reduces bias and may increase precision. Tangen and Koch in 1999 introduced a nonparametric covariate-adjusted method to estimate survival rates for certain given time intervals, relying only on the assumption that there is no association between covariates and treatment groups in a randomized clinical trial. We extended this nonparametric covariate-adjusted method to develop a new model to estimate the survival rates for treatment groups at any time point when an event occurs. Simulation studies are conducted to investigate the model performance. This model is illustrated with an oncology clinical trial example.

Comparing Strategies in Detecting Nonproportional Hazards

✱Yanqiong Zhang, Merck & Co., Inc, East Windsor, NJ 08520, *zhanyanq@yahoo.com*; Amarjot Kaur, Merck & Co., Inc

Key Words: nonproportional hazards, Schoenfeld score test

Testing of proportional hazard assumption in time-to-event analysis using Cox model is important for both substantive and methodological reasons. The proportional hazard assumption entails the relative hazards of any two levels of a covariate such as treatment group is constant over time. There are many methods in literature for detecting the nonproportional hazards assumption that entail formal testing and/or corresponding graphical approaches. The focus of this presentation is to examine different techniques, such as test of function of time-by-treatment interaction; partitioning time axis, and Schoenfeld score tests. Practically, there are no systematic comparisons available for different strategies under different scenario related to proportional hazards violations. An exploratory investigation comparing power of various strategies under many violation scenarios will be presented.

Using Marginal Structural Model To Adjust for Post-Study Treatment

✱Yanping Wang, Eli Lilly and Company, 14281 Chariots Whisper Dr., Westfield, IN 46074, *wang_yanping@lilly.com*; Shengyan Hong, Eli Lilly and Company; Min Zhang, North Carolina State University

Key Words: cancer clinical trials, causal inference, survival analysis

Comparison of survival between treatment arms in randomized cancer clinical trials is often complicated by post-study treatment (PST). Recently, the marginal structural model (MSM) proposed by Robins has been used by several authors to adjust for the effect of PST so the estimated treatment arm effects have the desired causal interpretation. Though theoretically appealing, this approach needs to be further evaluated in this setting. We conducted simulations to evaluate the performance of this approach under different situations.

Estimating a Treatment Effect in Oncology Clinical Trials with Correcting for Crossover

✱Qiang Xu, Columbia University, Department of Biostatistics, New York, NY , *qx2001@columbia.edu*; Xin Huang, Pfizer Global Research and Development; Vanessa Tassell, Pfizer Global Research and Development; Jim Li, Pfizer Global Research and Development; Randy Allred, Pfizer Global Research and Development

Key Words: survival analysis, crossover, oncology, clinical trial

Overall survival (OS) is recognized as the gold standard for clinical benefit in oncology clinical trials. However, the formal demonstration of the underlying survival benefit of a treatment can be diluted or confounded by effective later treatment. Although oncology trials have the objective of comparing the experimental arm with the control arm for OS, statistical analyses of OS by ITT treatment groups based on conventional approaches are biased towards the null because of the crossover. We evaluated two approaches, inverse probability of censoring weighted method and rank preserving structural failure time model, with data obtained from a randomized placebo-controlled phase 3 oncology trail. The results from the aforementioned approaches are similar, and both indicate that the experimental treatment could significantly improve the median OS compared to the placebo.

Comparison of Cox Regression Model, Stratified Cox Model, and Logrank Test on Statistical Power in Clinical Trials with Prognostic Factors

Jianliang Zhang, MedImmune, Inc.; ✱Iksung Cho, MedImmune, Inc., One MedImmune Way, Gaithersburg, MD 20878, *choi@medimmune.com*

Key Words: Cox regression, logrank test, stratified analysis, statistical power, prognostic factor

In stratified randomized clinical trials with survival endpoint, data analysis could utilize either Cox regression model, where prognostic factors (PF) are included as covariates, Cox model stratified by PFs, or stratified logrank test. Simulation studies to investigate these approaches with respect to power showed following. (1) "Power loss" occurs when PFs are not adjusted for or stratified by. (2) Power is well recovered in Cox regression with PF adjustments. However, recovery of power in the stratified analyses still does not bring the power back to the designed level. Moreover, stratifying by too many PF factors may penalize power. These findings come from interplay between biases in estimated SEs and estimated treatment effects. These results suggest the commonly used stratified analysis methods may result in loss of power for survival endpoint studies in heterogeneous population.

Time-to-Event Analysis Considerations in the Medical Device Industry

✱Helen M. Chmiel, Zimmer Corporation, 1800 Center St., Warsaw, IN 46580, *hmchmiel@aol.com*

Key Words: Time to event analysis, Survival analysis, Proportional hazards, Cox regression, Time dependent covariates

In many multicenter, randomized, controlled clinical trials in the device industry, a primary or secondary endpoint, occurring during the course of the clinical trial, may be best analyzed with time to event modeling techniques. Further there may be an interest in long term follow-up as confirmatory evidence of the success for a particular measure. Often the analysis can be enhanced by consideration of covariates that may influence results. Correctly adding in covariates to the analysis as well as correctly assessing analysis assumptions is critical for proper interpretation of treatment effect in this setting. This investigation considers different model building strategies using the Cox regression model which include assessing fitted Proportional Hazards assumptions, methods of selecting model covariates, and utility of time-dependent covariates for this analysis.

489 Stochastic Approximation Monte Carlo ▲

Section on Bayesian Statistical Science, Section on Statistical Computing, IMS

Thursday, August 7, 10:30 a.m.–12:20 p.m.

An Overview of Stochastic Approximation Monte Carlo Algorithms

✱Faming Liang, Texas A&M University, Department of Statistics, College Station, TX 77843, *fliang@stat.tamu.edu*

Key Words: Stochastic Approximation, Markov Chain Monte Carlo, Convergence, Trajectory Averaging, Asymptotic Normality, Self-Adjusting Mechanism

The stochastic approximation Monte Carlo (SAMC) algorithm has been recently proposed in the literature as a general Monte Carlo and optimization algorithm. A remarkable feature of the algorithm is that it avoids the local-trap problem suffered by conventional MCMC algorithms; the algorithm can self-adjust the acceptance rate of the local Metropolis-Hastings moves such that each subregion of the sample space can be sampled with a desired frequency. In this talk, we will provide an overview of the algorithm from both the theoretical and practical aspects. Applications of the algorithm to a variety of computational problems, e.g., model selection, marginal density estimation, machine learning, phylogeny estimation, and global optimization, will be discussed.

The Sample Metropolis-Hastings Algorithm

Chuanhai Liu, Purdue University; ✱Andrew Lewandowski, Purdue University, Department of Statistics, 250 N. University Street, West Lafayette, IN 47907-2066, *alewand@stat.purdue.edu*

Key Words: Bayesian argument, Deconvolution, Dempster-Shafer Theory, Population Monte Carlo, Stochastic Approximation

The Metropolis-Hastings (MH) algorithm is a powerful tool used to derive most Markov chain Monte Carlo (MCMC) sampling schemes for Bayesian computation. In the past decade, researchers have introduced modifications to the MH algorithm such as Population Monte Carlo (PMC) and Stochastic Approximation Monte Carlo (SAMC) in an attempt to effectively use information from past samples. In the tradition of these methods, the Sample Metropolis-Hastings (SMH) algorithm is a MH-based algorithm which creates updates based on a stored sample of values. Examples and theoretical properties are discussed, and the SMH algorithm is compared to similar methods, such as MH, PMC, and SAMC.

Asymptotics of the Wang-Landau Algorithm

✱Yves F. Atchade, The University of Michigan, Dept of Statistics, 1085 South University, Ann Arbor, MI 48109, *yvesa@umich.edu*

Key Words: Monte Carlo methods, Adaptive Monte Carlo, Wang-Landau algorithm, Stochastic approximation

The Wang-Landau algorithm is a Monte Carlo algorithm that has generated much interest in the physics literature due to some spectacular performances. This talk will discuss the consistency of the algorithm and how it can be adapted to sampling problems of more interest in statistics.

490 Advanced Modeling in Remote Sensing of the Inner Earth ●▲

American Geophysical Union, Section on Physical and Engineering Sciences

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Statistical Issues in Imaging the Inner Earth

✱Ping Ma, University of Illinois, 725 S Wright St, Champaign, IL 61820-5710, *pingma@uiuc.edu*; Luis Tenorio, Colorado School of Mines; Maarten de Hoop, Purdue University; Ping Wang, Massachusetts Institute of Technology; Robert van der Hilst, Massachusetts Institute of Technology

Key Words: inverse problem, core-mantle boundary, statistical imaging

At a depth of ~2890 km, the core-mantle boundary (CMB) separates turbulent flow of liquid metals in the outer core from slowly convecting, highly viscous mantle silicates. The CMB marks the most dramatic change in dynamic processes and material properties in our planet, and accurate images of the structure at or near the CMB—over large areas—are crucially important for our understanding of present day geodynamical processes and the thermochemical structure and history of the mantle and mantle-core system. In addition to mapping the CMB we need to know if other structures exist directly above or below it, what they look like, and what they mean. Detection, imaging, characterization, and understanding of structure in this remote region have been—and are likely to remain—a frontier in cross-disciplinary geophysics research.

Comparing Layer Transitions via Regularity Estimates

✱Luis Tenorio, Colorado School of Mines, The Center for Wave Phenomena, Golden, CO 80401, *ltenorio@mines.edu*

Regularity of a function at a point can be measured using what are known as regularity exponents, of which there is a wide variety (e.g., H”older exponent, chirp exponent, oscillating exponent, weak exponent, and p exponent). We present the basic definitions, the challenges in their estimation, and some preliminary results.

Sequential Estimation of High-Dimensional Space-Time Models

✱Jonathan Stroud, University of Pennsylvania, Philadelphia, PA 19104-6302, *stroud@wharton.upenn.edu*

Kalman filter methods for real-time assimilation of observations and dynamical systems typically assume knowledge of the system parameters. However, relatively little work has been done on extending state estimation procedures to include parameter estimation. Here, in the context of the ensemble Kalman filter, a Monte Carlo-based algorithm is proposed for sequential estimation of the states and model parameters. A Bayesian approach is adopted that yields analytical updating of the parameter distribution and provides samples from the posterior distribution of the states and parameters. The proposed assimilation algorithm extends standard ensemble methods, including perturbed observations, and serial and square root assimilation schemes. The method is illustrated on the Lorenz 40-variable system and a real example involving assimilation of satellite reflectance images.

491 Nutrition Surveys: Challenges in Design and Analysis ●

SSC, Section on Survey Research Methods, Social Statistics Section, Section on Statistics in Epidemiology

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Under-Reporting of Energy Intake in the Canadian Community Health Survey

✱Didier Garriguet, Statistics Canada, RH Coats Building 24 K, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *didier.garriguet@statcan.ca*

Key Words: energy underreporting, energy intake, nutrition, energy expenditure

Under-reporting is a key issue when dealing with nutrition surveys. This paper estimates under-reporting of energy in the Canadian Community Health Survey (CCHS), identifies characteristics of under-reporters and assesses the impact of under-reporting on data analysis. Total energy expenditure based on equations from the Institute of Medicine is compared to energy intake. Confidence intervals inspired by the Goldberg cutoffs are used to identify under-reporters. In CCHS, under-reporting is estimated at 9.6%. Age, sex, being overweight or obese and physical activity are associated with under-reporting. Identifying under-reporters results in a positive association between energy intake and being obese.

Analysis of Survey-Based Usual Intake Nutrition Data: The Issue of Within-Person Variation

✱Kevin Dodd, National Cancer Institute, 6130 Executive Blvd, Rockville, MD 20892, *doddk@mail.nih.gov*; Joseph Goldman, U.S. Department of Agriculture

Key Words: 24-hour recall, usual intake, nonlinear mixed model, balanced repeated replication

Statistical methods to estimate usual (i.e., long-run average) daily intake of a food or nutrient are required to properly conduct dietary surveillance. The dietary assessment portion of the NHANES includes up to two 24-hour recalls (24HR) per person. Day-to-day (within-person) variability in diet ensures that 24HRs measure usual intake with substantial error. For dietary components not consumed daily, 24HR reports may be zero, even for individuals with positive usual intake. Distributions of measured and usual intake are often very skewed. We present a nonlinear mixed model for usual intake and demonstrate its use in estimating the population distributions of usual intake for major food groups and nutrients. The balanced repeated replication method was used to approximate the standard errors of estimated quantiles of usual intake.

National Nutrition Data: Methodological and Analytic Experiences of NHANES

✱Clifford L. Johnson, Centers for Disease Control and Prevention, 3311 Toledo Road, Hyatsville, MD 20782, *CLJohnson@cdc.gov*; Leyla Mohadjer, Westat; Lester R. Curtin, Centers for Disease Control and Prevention

Key Words: dietary methods, NHANES, nutrition

This presentation provides a history of the evolution of the methods for the collection and analysis of nutrition data at NCHS/CDC. The NHANES Survey began collecting nutrition data in 1971. The nearly 40 years of NHANES experience is summarized in terms of the many changes in mode (paper questionnaires, CAPI, and telephone), methods (repeated 24 hour recalls and food frequency), survey participation rates, and the data base for processing the nutrition information. In addition, the survey design and the statistical methods for analyzing the data have changed over time. NHANES has moved from a 6-year periodic survey to an annual survey with data released in two year cycles. The current design is recognized to have analytic issues related to seasonality and small number of PSUs. The impact of these changes will be discussed in terms of their impact on time trends for nutrition data.

492 Radiation Exposure Effects Research: Moving Forward After 60 Years of Following A-Bomb Survivors ●

Section on Statistics in Epidemiology, Section on Statisticians in Defense and National Security, WNAR, Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Radiation-Related Cancer for Childhood Exposures

✱Donald Pierce, Oregon Health & Science University, 97239-3098, *pierce.don.a@gmail.com*

Key Words: radiation, cancer, exposure-age

Much of the emerging information in A-bomb survivor studies necessarily pertains to those exposed as children. Age-time patterns of their radiation-related cancer rates may shed light on basic mechanisms of carcinogenesis. Radiation cancer risks persist for all of lifetime, although with decreasing relative risk since the added mutations are a diminishing fraction of somatic accumulations. Effects of exposure-age on excess cancer risks have long been of primary interest. However, this involves difficult issues, namely a complicated sort of confounding with secular trends in baseline cancer rates. Resolution of this hinges on whether factors causing the secular trends act multiplicatively or additively with radiation, which will depend on the cancer type. On another matter, it appears that there may be greater upward curvature in the dose response for those exposed as children.

Application of Causal Modeling on Radiation, Inflammation, and Selected Radiation/Inflammation-Induced Health Outcomes

✱Tatsuyuki Kakuma, Kurume University, 67 Asahi-Mach, The Biostatistical Center, Kurume, 830-0011 Japan, *tkakuma@med.kurume-u.ac.jp*; Yuko Araki, Kurume University; Wan-Ling Hsu, The Radiation Effects Research Foundation; Eiji Nakashima, The Radiation Effects Research Foundation; Kazuo Neriishi, Radiation Effects Research Foundation; Ritsu Sakata, The Radiation Effects Research Foundation; Norman P. Ross, Radiation Effects Research Foundation

Key Words: Low-Grade Inflammation, Joint Model, Radiation-Induced Diseases

Atomic bomb survivors have persistently increased risk of cancer and non-cancer diseases. The mechanisms of how radiation exposure causes these undesirable health conditions are quite complex. One hypothesis is that radiation exposure results in chronic low-grade inflammation, which in turn may cause elevated risk for some inflammation-related and radiation-induced diseases such as cataract. Inflammation is usually indicated by various biomarkers, and we hypothesized a causal pathway involving radiation, inflammation, inflammation-induced covariates and other risk factors among the atomic bomb survivors in the Adults Health Study. Association among these variables were examined by a joint model where onset of cataract was treated as survival data, latent inflammation, inflammation-induced covariates, and other risk factors including radiation dose were modeled with a structural model.

Bayesian Risk Projection for A-Bomb Survivor Cohort

✱Kyoji Furukawa, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku,, Hiroshima City, International 732-0815 Japan, *furukawa@rerf.or.jp*; John Cologne, Radiation Effects Research Foundation; Yukiko Shimizu, Radiation Effects Research Foundation

Key Words: Age-period-cohort models, Cancer incidence and mortality, Risk assessment, Bayesian prediction

The number of cases causally related to the exposure under study, or excess cases, is an important determinant of statistical power for assessing aspects of risk such as age-time trends and susceptible subgroups. In determining how large a population to study or how long to follow a study population to accumulate sufficient excess cases, it is necessary to predict future risk. In this study, focusing on models involving excess risk with effect modification, we propose a method for predicting the expected magnitude of numbers of excess cases and assess the uncertainty in those predictions. We do this by extending Bayesian age-period-cohort models for rate projection to include excess risk and modification by age at exposure. The method is illustrated using the follow-up study of Japanese Atomic-bomb Survivors, a primary base for determining long-term health effects of radiation exposure.

Estimation of Radiation Dose from Biological Manifestations and Imperfect Measures of Physical Determinants

✱Randolph L. Carter, University at Buffalo, 249 Farber Hall, 3435 Main Street, Bldg 26, Buffalo, NY 14214-3000, *rcarter@buffalo.edu*; Harry Cullings, Radiation Effects Research Foundation; John Cologne, Radiation Effects Research Foundation; Sachiyo Funamoto, Radiation Effects Research Foundation; Yoichiro Kusunoki, Radiation Effects Research Foundation; Kazuo Neriishi, Radiation Effects Research Foundation; Thomas Seed, Radiation Effects Research Foundation; Austin Miller, University at Buffalo; Carmen D. Tekwe, University at Buffalo; Nori Nakamura, Radiation Effects Research Foundation; Daniel Stram, University of Southern California; Norman P. Ross, Radiation Effects Research Foundation

Key Words: measurement error, Berkson error, instrumental variable estimation, estimation bias, biodosimetry

Current dosimetry system (DS02) estimates of atomic bomb radiation dose are based on physical determinants. Two distinct types of errors exist in DS02 values, with statistical characteristics of measurement error and Berkson error. With measurement error alone in a linear model, dose effects are not identified without additional information. We use biological manifestations of dose as instrumental variables to identify parameters. With Berkson error alone, the model is identified. With both types of error present, a known/estimated measurement error variance suffices to identify the model, but an instrumental variable alone does not. An instrumental variable plus estimated Berkson error variance suffices. We propose methods for estimating model parameters and for combining information in biological dosimeters with that in physical dosimeters to improve estimates of radiation dose.

493 Global Maximization in EM-Type Algorithms

Section on Statistical Computing, Section on Bayesian Statistical Science

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Adaptive Data Weighting Strategies for Locating the Global Maximum in EM-Type Algorithms

✱Ravi Varadhan, Johns Hopkins University, 2024 E. Monument Street, School of Medicine, Suite 2-700, Center on Aging and Health, Baltimore, MD 21205, *rvaradhan@jhmi.edu*

Key Words: local maxima, latent class models, finite mixtures, EM acceleration, global maximization, squared iterative methods

We explore and evaluate some data weighting strategies that adaptively reweigh the data to facilitate convergence to a better local maximum than that achieved by the standard EM algorithm. The adaptive data weighting strategies exploit the special characteristics of the EM algorithm. We will also show how they can be combined with the SQUAREM acceleration schemes discussed in Varadhan and Roland (*Scandinavian Journal of Statistics*, 2007) to obtain fast converging iterative schemes. We will evaluate the effectiveness of these strategies in two problems: a simulation example involving a finite, Gaussian mixture; and a real-world problem involving latent class modeling of the profile of multiple biomarkers in a geriatric syndrome.

Global Optimization with Model Reference Adaptive Search and Expectation-Maximization

✱Jeffrey Heath, Centre College, 600 West Walnut St., Danville, KY 40422, *jeffrey.heath@centre.edu*; Michael Fu, University of Maryland; Wolfgang Jank, University of Maryland

Key Words: EM Algorithm, Model Reference Adaptive Search, Global Optimization, Mixture Models

It is well-known that the likelihood function of Gaussian mixture models can have many local, suboptimal maxima. While the Expectation-Maximization (EM) algorithm is the standard tool for estimating mixture-model parameters, it is easily trapped into such local maxima. We propose a systematic way of estimating mixture-model parameters based on the global optimization method Model Reference Adaptive Search (MRAS). One of the advantages of MRAS is that global convergence can be proved rigorously. We adapt MRAS to the Gaussian mixture model, and provide a theoretical proof of global convergence to the optimal solution of the likelihood function. We combine the updating procedure of MRAS with that of EM to construct the MRAS-EM algorithm for Gaussian mixtures. We provide numerical experiments which illustrate the performance of the MRAS algorithm relative to the EM algorithm.

A Probabilistic Analysis of EM for Mixtures of Spherical Gaussians

✱Sanjoy Dasgupta, University of California, San Diego, 9500 Gilman Drive #0404, La Jolla, CA 92109, *dasgupta@cs.ucsd.edu*

Key Words: expectation maximization, clustering, mixtures of Gaussians, probabilistic analysis

The EM algorithm for fitting Gaussian mixture models is one of the most widely-used clustering methods. Yet, surprisingly little is known about its behavior. Under what conditions does it converge to a good local optimum? What are good ways to initialize it? Such questions are difficult to answer with the mathematical tools that have traditionally been applied to EM. I will describe an alternative way of analyzing EM: by a probabilistic analysis. This reveals, first of all, that many common methods of initializing EM produce highly suboptimal results even in ideal clustering scenarios. On the other hand, I'll show that a particular variant of EM will provably recover a near-optimal clustering, provided that the clusters are adequately separated and that their distributions are weakly Gaussian.

494 Imaging Biomarkers in Oncology ●▲

Biopharmaceutical Section, WNAR, Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Practical Considerations in Exploratory Imaging

✱Haren Rupani, Novartis Pharmaceuticals, 1 Health Plaza, East hanover, NJ 07936, *haren.rupani@novartis.com*

Key Words: Exploratory imaging, qualification, biomarkers, Oncology

In light of the current availability of functional and molecular imaging in addition to the traditional anatomic imaging modality for Oncology drug development, an exploratory imaging strategy is used to develop potentially useful biomarkers. The qualification of these biomarkers is actively being pursued through many public-private partnerships. The goal of this talk is to share the Novartis experience in using exploratory imaging in oncologic drug development.

Imaging Modalities Can Predict Therapy Outcomes: Can They Be Used To Define Endpoints in Clinical Trials?

✱Constantine Gatsonis, Brown University, Center for Statistical Sciences, Box G-S121-7, Providence, RI 02912, *gatsonis@stat.brown.edu*

Key Words: biomarkers, imaging, clinical trials, endpoints

Diagnostic imaging is already a routine part of the assessment of endpoints in clinical studies. For example, RECIST criteria are used broadly in oncology to define response to therapy. However, a new range of possibilities has been opened by the advent of functional and molecular imaging techniques, which have the potential to predict the outcome of therapeutic interventions very early. In this presentation we will examine the promise and the limitations of such imaging based biomarkers to serve as endpoints for clinical trials and to direct clinical practice.

Evaluation of Novel Imaging Agents in the Context of Treatment Effects

✱Lori E. Dodd, National Cancer Institute, *doddl@mail.nih.gov*

Key Words: diagnostic accuracy, imaging, biomarker

Positive and negative predictive values describe a markers predictive ability, while sensitivity and specificity describe its discriminating ability. The first Prentice criterion would be a nice condition for a biomarker to satisfy because then predictive values do not depend on a covariate (such as treatment) and the biomarker captures all relevant information about the clinical state of interest. A similar condition can be defined for sensitivity and specificity which states that these measures do not depend on a covariate. This condition is desirable because it allows sensitivity and specificity from one treatment setting to be applied to a different setting. We demonstrate, however, that the Prentice condition and EDA are incompatible. We discuss the implications of these results in the evaluation of imaging agents. Indeed, alternatives to sensitivity/specificity may be desirable

495 Statistical Analysis of Childhood Precursors of Adult Disease

ENAR, Section on Statistics in Epidemiology, WNAR, Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

An Overview of Longitudinal Studies Useful to Investigators Studying Childhood Determinants of Adult Disease

Elizabeth Goodman, Tufts-New England Medical Center/The Floating Hospital for Children; ✱Terry T.K. Huang, National Institutes of Health, Eunice Kennedy Shriver National Institute of Child Health and Human Development, 6100 Executive Boulevard, 4B11, MSC 7510, Bethesda, MD 20892-7510, *huangter@mail.nih.gov*

Key Words: risk factors, longitudinal models, childhood

This presentation will describe several longitudinal studies which are available to investigators interested in studying childhood precursors of adult disease. These databases include the National Longitudinal Study of Youth, the National Longitudinal Study of Adolescent Health, the Bogalusa Heart Study, the National Growth and Health Study, and the Fels Longitudinal Study. All of these studies include familial and child specific data on health. Some also have collected genetic material and others have contextual variables such as geocoded environmental data. For each study, an overview of the study design, variables collected, and time course for data collection will be reviewed. In addition, commonalities and differences between studies will be highlighted.

Analytic Approaches: Epidemiological, Statistical, and Genetic

✱Christine M. Schubert, Virginia Commonwealth University, Theatre Row, Room 3084, 730 East Broad Street, SOM, P.O. Box 980032, Richmond, VA 23298-0032, *cmschubert@vcu.edu*

Key Words: longitudinal modeling, serial data, ROC, methods, genetic modeling

The study of childhood precursors to adult disease requires a study design following individuals over time. The design of the Fels Longitudinal Study offers the opportunity to apply a variety of analytical methods to describe and predict adulthood disease and offers insight into phenotypes which, if intervened, may decrease chances of adult disease onset. This presentation demonstrates various analytical approaches applied to the Fels Longitudinal data to address prediction of adult disease outcomes. Analytical methods to be addressed include, long-term serial data regression modeling (mixed effect modeling), logistic and ROC curve methodology, and genetic algorithms to analyze familial effects. Results from analyses of the childhood Fels Longitudinal Study will be demonstrated for the prediction of adulthood diseases such as metabolic syndrome, hypertension and CVD.

The Fels Longitudinal Study of Growth and Human Development

✱William C. Chumlea, Wright State University, Department of Community Health Boonshoft School of Medicine, 3171 Research Blvd, Dayton, OH 45420, *cameron.chumlea@wright.edu*; Audrey Choh, Wright State University; Miryoung Lee, Wright State University; Bradford Towne, Wright State University; Dana Duren, Wright State University; Stefan Czerwinski, Wright State University

Key Words: Fels, Serial, body composition, lifespan, quality control, risk factors

The Fels Longitudinal Study is the longitudinal study of growth development and aging in the world. It consists of extensive serial data collected across the lifespan for 1500 related individuals. Participants are unselected based on health status or any specific disease or related trait. Each participant is followed from birth until death regardless of health. At visits age-appropriate data are collected that allow analysis of changes in growth body composition and risk factors across the lifespan and of relationships and changes in relationships between growth and risk factors within and among related individuals over time from childhood into old age such as between CVD or diabetes and changes in growth and other risk factors that occurred earlier in childhood. Quality control is a high priority; data are checked regularly for errors along with data collection standardization.

496 Statistical Methods and Applications of Social Network Analysis ●

Section on Health Policy Statistics, Section on Statisticians in Defense and National Security, Social Statistics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Spatial Process Model for Social Network Analysis

✱Crystal Linkletter, Brown University, Box G-S121-7, 121 South Main St., Providence, RI 02912, *cdlinkle@stat.brown.edu*

Key Words: social networks, Bayesian inference, spatial process

There has been a recent increase in the use of network models for representing interactions and structure in many complex systems. Here we propose using a latent spatial process model for the statistical analysis of networks, emphasizing applications to social networks. This approach smoothes the relationship between connections and covariates in a sample network with relatively few parameters, allowing probabilities of connection for a population to be inferred. One important application of such a predictive network model is to improve understanding of how an infectious disease might spread in a population, since local network topology can have a significant impact on contact-based processes.

Recent Statistical Models for Network Science

✱Stanley Wasserman, Indiana University, 309 North Park Avenue, Statistics House 202, Bloomington, IN 47408, *stanwass@indiana.edu*

The purpose of this short exposition is to discuss the developments in statistical models for networks that have occurred since the publication of the statistical chapters (8, 9, 10, and 11) of Carrington, Scott, and Wasserman (2005, Cambridge University Press).

Factor Models for Multivariate Relational Data

✱Peter Hoff, University of Washington, Department of Statistics, Seattle, WA 98195-4322, *hoff@stat.washington.edu*

Key Words: social network, stiefel manifold, multivariate analysis, random matrix

Much of network and relational data is multivariate: Between-node links are often of multiple types and measured under a variety of conditions and time-points. Node-level data are frequently gathered as well. Often of interest are the correlations between different network measurements and their relationship to node-level attributes. Unfortunately, most of the tools for analysis of such data are univariate. Regression models of one relational variable as a function of other relational and nodal variables ignore the dependencies of the "independent" variables. This talk describes a statistical model for the joint distribution of dyad and node-level ordinal variables. The model is based on an extension of the idea of factor analysis: Each node has a vector of unobserved latent factors, and measurements (both dyadic and node-specific) involving a node are functions of it factors.

497 Advances in Variable Selection ●

IMS, SSC

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Variable Selection via a Bayesian Ensemble

✱Hugh Chipman, Acadia University, Deptartment of Mathematics and Statistics, Wolfville, NS B4P1Z1 Canada, *hugh.chipman@acadiau.ca*; Edward George, The Wharton School, University of Pennsylvania; Robert McCulloch, The University of Chicago

Key Words: ensemble, MCMC, statistical learning, tree

Bayesian methods provide an attractive and comprehensive approach to learning ensemble models such as forests of trees. We consider a particular case: the flexible and fast Bayesian Additive Regression Trees (BART). BART can be used to screen for relevant predictors, providing an essentially nonparametric approach to variable selection. As the BART algorithm runs, different potential predictors enter the sum-of-trees model with different frequencies. Those that enter rarely or not at all are candidates for elimination, and those that enter frequently are candidates for inclusion. By varying the size of the sum-of-trees model, BART can identify subsets of predictors containing the strongest predictive information. BART also provides an omnibus test: the absence of any relationship between y and the predictors is indicated when BART posterior intervals for the response reveal no signal.

Alpha Investing: A New Multiple Hypothesis Testing Procedure That Controls mFDR

✱Dean Foster, University of Pennsylvania, Philadelphia, PA 19103, *dean@foster.net*; Robert A. Stine, University of Pennsylvania

Key Words: regression, variable selection, stepwise, data mining

We propose alpha investing, an adaptive, sequential methodology for testing multiple hypotheses. Alpha investing encompasses a large family of rules for testing multiple hypotheses, all having the Property that the mFDR is controlled. mFDR, which is the ratio of the expected number of false rejections to the expected number of rejections, is a weaker criterion than the FDR, which is the expected value of the ratio. We partially compensate for this weakness by showing that alpha-investing rules have a stronger martingale property that offers control of the procedure beyond mFDR. Alpha-investing rules mimic alpha-spending rules used in sequential trials, but possess a key difference. When a test rejects a null hypothesis, alpha-investing rules earn additional probability toward subsequent tests.

Real-Time Prediction Under Model Uncertainty via Dynamic Model Averaging

✱Adrian E. Raftery, University of Washington, Department of Statistics, Box 354320, Seattle, WA 98195-4320, *raftery@u.washington.edu*; Miroslav Karny, Institute of Information Theory and Automation; Josef Andrysek, Institute of Information Theory and Automation; Pavel Ettler, Compureg

Key Words: Model uncertainty, Markov chain, Bayesian model averaging, automatic control, state space model, forgetting

We consider the problem of real-time prediction when it is uncertain what the best prediction model is. We develop a method called Dynamic Model Averaging (DMA) in which a state space model for the parameters of each model is combined with a Markov chain model for the correct model, allowing the correct model to vary over time. The state space and Markov chain models are both specified parsimoniously in terms of forgetting. The method is applied to predicting the output of a cold rolling mill, where the output is measured with a time delay. When only a small number of physically-based models were considered and one was clearly best, the method quickly converged to the best model, and the cost of model uncertainty was small. When model uncertainty and the number of models considered were large, our method ensured that the penalty for model uncertainty was small.

498 New Developments on Analysis of Correlated Data ●

WNAR, Biopharmaceutical Section, Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Variable Selection for High-Dimensional Correlated Data

✱Xihong Lin, Harvard University, Department of Biostatistics, Harvard SPH, 655 Huntington Ave, Boston, MA 02115, *xlin@hsph.harvard.edu*

Key Words: Variable selection, Penalized Likelihood, Correlated data, Sparseness, Asymptotics, High-dimensional data

We propose several variable selection methods for correlated data, such as longitudinal and clustered data, based on regularized likelihood functions. We study the theoretical properties of these methods, and compare their theoretical performance. Their finite sample performance is evaluated using simulation studies and illustrated using a data example.

Forecasting Time Series of Inhomogeneous Poisson Processes

✱Jianhua Huang, Texas A&M University, Department of Statistics, College Station, TX 77843, *jianhua@stat.tamu.edu*; Haipeng Shen, The University of North Carolina at Chapel Hill

Key Words: dimension reduction, factor model, penalized likelihood, queueing systems, service engineering, vector time series

We consider forecasting the latent and uncertain rate profiles of a time series of inhomogeneous Poisson processes. The work is motivated by operation management of queuing systems such as call centers. Our forecasting approach utilizes dimension reduction through a factor analysis of Poisson variables, followed by time series modeling of factor score series. Time series forecasts of factor scores are combined with factor loadings to yield forecasts of future Poisson rate profiles. Penalized Poisson regressions on factor loadings guided by time series forecasts of factor scores are used to generate dynamic within-process rate updating. Methods are also developed to obtain distributional forecasts. Our methods are illustrated using simulation and real data. In particular ,we show how forecasting and dynamic updating of call arrival rates can affect the accuracy of call center staffing.

Generalized Varying Coefficient Models for Longitudinal Data

✱Damla Senturk, The Pennsylvania State University, 411 Thomas Building, University Park, PA 16801, *dsenturk@stat.psu.edu*; Hans G. Müller, University of California, Davis

Key Words: Linear regression, Measurement error model, Prediction, Smoothing, Two-step procedure

We propose a generalization of the varying coefficient model for longitudinal data to cases where not only current but also recent past values of the predictor process affect current response. More precisely, the targeted regression coefficient functions of the proposed model have sliding window supports around current time t. A variant of a recently proposed two-step estimation method for varying coefficient models is proposed for estimation in the context of these generalized varying coefficient models, and is found to lead to improvements, especially for the case of additive measurement errors in both response and predictors. Asymptotic distributions of the proposed estimators are derived, and the model is applied to the problem of predicting protein concentrations in a longitudinal study. Simulation studies demonstrate the efficacy of the proposed estimation procedure.

Partial Consistency in Mixed Models

✱Tao Huang, University of Virginia, Department of Statistics, Charlottesville, VA 22904, *th8e@virginia.edu*; Heng Peng, Hong Kong Baptist University

Key Words: Partial Consistency, Mixed Models, Correlated Data

The concept of partial consistency is extended to mixed models in the analysis of correlated data. A two-step method is proposed to estimate the random and fixed effects. Comparing to classical estimation methods, the proposed two-step estimation method relaxes assumptions on random effects and alleviates the complexity of computation. The asymptotic properties, especially the efficiency and robustness, of the proposed two-step method will be addressed, and the applicability and effectiveness of the proposed two-step method will be illustrated through simulation studies and real data analyses.

499 Teaching Introductory Statistics Online ▲

Section on Statistical Education, Section on Teaching Statistics in the Health Sciences

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Teaching Introductory Statistics Online

✱Felicity B. Enders, Mayo Clinic, Department of Health Sciences Research, 200 First Street, SW, Division of Biostatistics, Rochester, MN 55905, *Enders.Felicity@mayo.edu*; ✱Engin Sungur, University of Minnesota, Morris, 600 E. 4th Street, Morris, MN 56267, *sungurea@morris.umn.edu*; ✱John McGready, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Room E3543, Baltimore, MD 21205, *jmcgread@jhsph.edu*; ✱Gail Tudor, Husson College, One College Circle, Bangor, ME 04401, *tudorg@husson.edu*

Key Words: statistics education, online education, evaluation

Recent technological innovations, together with a push to teach students at a distance, have coalesced in a desire at many institutions to put statistics courses online. However, not all online courses are created equal. This session will include three half hour talks followed by panel discussion and questions. Topics include 1) creative content options in an online course (Dr. Sungur), 2) software tools for developing a strong online course (Dr. Boyd Enders), and 3) evaluation of learning in an online environment (Dr. McGready). All three talks will be based on real-life interactive introductory statistics courses which are taught online. Dr. Tudor, the discussant, will lead panel discussions on each topic.

500 Hypothesis Testing in Neuroimaging ●

Biometrics Section, Biopharmaceutical Section,
WNAR
Thursday, August 7, 10:30 a.m.–12:20 p.m.

Type I Error Rate with Simple Group fMRI Models

✱Jeanette A. Mumford, University of California, Los Angeles, 1946 Overland Ave, Apt 206, Los Angeles, CA 90025, *mumford@ucla.edu*; Thomas E. Nichols, GlaxoSmithKline Clinical Imaging Centre

Key Words: fMRI, group modeling, Ordinary Least Squares, type I error rate

Group modeling of functional magnetic resonance imaging (fMRI) data is done according to one of two broad approaches: Ordinary Least Squares (OLS) analysis of contrast images, and a mixed models (MM) approach, where optimal weighting is based on between- and within-subject variance estimates. Despite widespread use of OLS for group modeling of fMRI data, a careful evaluation of assumptions and limitations of the method has not been performed. Different authors have proposed MM methods for fMRI, but used narrow assumptions in their evaluations or only consider a limited class of studies. In particular, previous authors have focused on the detrimental effects heteroscedasticity has on estimator precision, overlooking the impact on the null distribution of the test statistic. This work focuses on the null distribution of the OLS statistic under different levels of heteroscedasticity.

Bootstrap Procedures for Testing in Neuroimaging Settings

✱Todd Ogden, Columbia University, Department of Biostatistics, 6th floor, 722 W. 168th St., New York, NY 10032, *to166@columbia.edu*; Chung Chang, Columbia University

Key Words: family-wise error, bootstrap, brain imaging

A common situation in neuroimaging studies involves testing for a difference in means in images taken from two or more groups. This is typically done for each voxel, and the family-wise error rate (FWER) may be controlled by imposing some stringent parametric assumptions on the spatial covariance structure of the data (Statistical Parametric Mapping; SPM). When group labels are exchangeable, the FWER may be controlled using a permutation test. For instances in which labels are not exchangeable, we propose nonparametric bootstrap techniques for controlling the FWER that maintain the spatial covariance structure.

Inference for Eigenvalues and Eigenvectors of Diffusion Tensors in Multisubject Diffusion Tensor Imaging Studies

✱Armin Schwartzman, Harvard School of Public Health, Biostatistics Dept., Dana-Farber Cancer Institute, 44 Binney Street, Mailstop LW 225, Boston, MA 02115, *armins@hsph.harvard.edu*; Robert F. Dougherty, Stanford University; Jonathan E. Taylor, Stanford University

Key Words: random matrices, symmetric matrices, positive definite matrices, manifold data, likelihood ratio test, matrix log

Diffusion tensor imaging (DTI) data differ from most imaging data in that values at each voxel are not scalars, but 3 ◊ 3 positive definite matrices, called diffusion tensors (DTs). Multi-subject DTI studies often aim to find local anatomical differences between two groups of subjects, where anatomical features at each voxel are encoded in the DT's eigenvalues (tissue type and health) and eigenvectors (spatial orientation of neural fibers). Assuming a Gaussian model based on the central limit theorem for symmetric matrices, we derive likelihood ratio tests for testing whether the means of two samples of DTs have the same eigenvalues and/or eigenvectors. The inference problems are non-standard in that the parameter sets are curved embedded submanifolds of the Euclidean space of symmetric matrices. The tests are illustrated in a voxelwise comparison of dyslexic vs. nondyslectic children.

Simultaneous Confidence Bands for the Coefficient Function in Functional Regression

✱Philip Reiss, New York University, 215 Lexington Ave., 16th floor, New York, NY 10016, *reissp01@med.nyu.edu*; Todd Ogden, Columbia University

Key Words: Functional principal component regression, Functional generalized linear model, Neuroimaging, Simultaneous confidence bands, Smoothing parameter selection

We extend functional principal component regression, a recently proposed spline-based method for regressing scalar outcomes on functional predictors, to generalized linear models with brain images as predictors. To test simultaneously for a zero regression coefficient at each of a high number of voxels, we propose an approximate simultaneous confidence interval for the coefficient image, which can be inverted to perform the desired test. The interval estimate is derived by bootstrapping cases, which ordinarily tends to result in undersmoothed function estimates; a novel approach to smoothing parameter selection is proposed to correct this tendency.

Intrinsic Regression Model for Positive-Definite Matrices with Applications to Diffusion Tensor Imaging

✱Yimei Li, The University of North Carolina at Chapel Hill, 3101 McGavran-Greenberg, CB#7420, Biostatistics, Chapel Hill, NC 27599, *liyimei@email.unc.edu*; Hongtu Zhu, The University of North Carolina at Chapel Hill; Yasheng Chen, The University of North Carolina at Chapel Hill; Joseph G. Ibrahim, The University of North Carolina at Chapel Hill; Weili Lin, The University of North Carolina at Chapel Hill

Key Words: Diffusion tensor, Differential geometry, Regression, Resampling method, Tangent space

Because the positive-definite matrices do not form a vector space, classical multivariate regression may undermine their association with covariates in Euclidean space. We propose a semiparametric regression model to solve this problem. The intrinsic mean of positive-definite matrices included as a special case. We develop an estimation procedure to calculate parameter estimates and establish their limiting distribution. We develop score statistics to test linear hypotheses of unknown parameters and a test procedure based on a resampling method to simultaneously assess the statistical significance of hypotheses across region of interest. Simulation studies shows the validity of our test procedure. We then apply our methods to the detection of statistical significance of diagnostic effect on the integrity of white matter in a diffusion tensor study of multiple sclerosis.

501 Design and Inferences in Multiregional Clinical Trials ●

Biopharmaceutical Section, Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Bridging Estimate of Treatment-by-Race Effect Using Balancing Scores

✱Kyoungah See, Eli Lilly and Company, 450 S. Meridian Street, Faris Campus II, Mail Drop Code 6132, Indianapolis, IN 46225, *seeky@lilly.com*; Ilya Lipkovich, Eli Lilly and Company

Key Words: Matching Scores, Size Adjusted, Bootstrap

The primary objective of most bridging studies in clinical development is to allow data from a secondary region (e.g., Japan) to be utilized in a trial with the data from a primary region (e.g., Western), based upon demonstrated parallel behavior of the two regions' respective dose-response curves. One of the main challenges of the bridging study is the difference in sample size and baseline covariates potentially affecting the study outcome. For example, the Western data in clinical trials on osteoporosis is often of much larger sample size than its Asian counterpart while the distribution of many important risk factors of fractures may differ substantially across the two regions.

Harmonization of Statistical Analyses and Reporting in a Global Environment: Some Experiences, Challenges, and Solutions

✱Melvin S. Munsaka, Takeda Global Research & Development, One Takeda Parkway, Deerfield, IL 60015, *mmunsaka@tgrd.com*

The conduct of global clinical trials presents many problems and challenges including logistical, analysis, data issues, and regulatory considerations. From a statistical perspective, it is important that the issues regarding analysis are harmonized among regions, especially if data and analyses from studies done in one region are to be used as part of a regulatory submission or to support a submission in another region. Some level of harmonization and cooperation in statistical activities and planning is necessary to allow for efficient integration of analyses for regulatory submissions as well as for quick and efficient assessment of efficacy and safety. In this talk, we will focus on the issues that surround statistical analyses and reporting in a global environment and discuss some approaches aimed at gaining global harmonization drawing on some experiences and examples.

Issues of Analysis and Interpretation Relating to Regions in Multiregional Clinical Trials

✱Paul Gallo, Novartis Pharmaceuticals, One Health Plaza, 419/2114, East Hanover, NJ 07936, *paul.gallo@novartis.com*

Key Words: interaction, subgroup analysis, heterogeneity, bridging study

There is much current interest in design and analysis issues for trials run in multiple geographic areas. This is enhanced when local health agencies desire to obtain not only overall proof of effect but also some assurance of effect within their domain. In trials designed conventionally to show an effect in a full population, definitive proof of effect in subsets will not be feasible to achieve with high power. We can expect to be left with some ambiguity about heterogeneity across region, as for other subgroups. We consider whether the factor 'region' is so unique as to merit its own analysis strategies relative to how other subgroups are typically handled. A challenge is that 'region' may at times be an artificial construct, confounded with other factors that are the true effect modifiers. We address how these issues can be reflected in trial design, analysis, and interpretation.

Issues in Conducting and Evaluating Multiregional Clinical Trials

✱Yuki Ando, Pharmaceuticals and Medical Devices Agency, Shin-Kasumigaseki Bldg., 3-3-2, Kasumigaseki, Chiyodaku, Tokyo, International 100-0013 Japan, *ando-yuki@pmda.go.jp*

Key Words: global drug development, multi-regional trial

Recently Japanese pharmaceutical companies have planned and started to conduct multiregional clinical trials. Such trials are planned not only as confirmatory trials but also as exploratory trials or dose response studies. We have to consider how the data from such trials will be a useful component of complete clinical data package for new drug application. In Japan, "Basic Principles on Global Clinical Trials" was issued in Sep. 2007, based on PMDA's experience of new drug review and clinical trial consultation. In this presentation, recent issues in conducting and evaluating multiregional clinical trials raised in Japan will be reviewed

502 Statistical Applications in Business

Business and Economics Statistics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

How To Help Businesses Make Sampling Decisions

✱Eric T. Falk, Ernst & Young LLP, 1101 New York Ave. NW, Washington, DC 20005, *eric.falk@ey.com*

Key Words: Statistical Sampling, Precision, IRS Stat Sampling Requirements

Businesses are increasingly using statistical sampling as a way to more efficiently quantify their business related activity. Currently businesses use statistical sampling to estimate meals and entertainment deductions, qualifying research expenditures and qualifying domestic manufacturing deductions to name a few. The Internal Revenue Service allows statistical sampling when there is a "facts and circumstances" type situation and the cost and time to research all records in the population is unreasonable. However, there are restrictions to the use of statistical sampling by the IRS and taxpayers are forced to use least advantages confidence interval bounds from statistical samples not meeting certain precision requirements. This paper will discuss analysis and advice that is provided to companies so that they can plan whether statistical sampling will help them make tax decisions.

Statistical Documentation Practices for Business Consulting

✱Susan Hinkins, National Opinion Research Center, 1122 South 5th Ave, Bozeman, MT 59715, *hinkins-susan@norc.org*; Edward Mulrow, National Opinion Research Center

Statisticians know that an important part of statistical practice is the documentation of the data and statistical methodology used. Statistical consulting in business often involves working with government regulatory agencies that have their own minimal reporting standards. These standards are not necessarily sufficient for the needs of statisticians. In particular, the consultant may need to come back to describe (or defend) a project two to three years after its completion. This paper will provide several examples of the authors' experience with such documentation requirements and rules, and discuss what practices are important to meet both the agency's requirements and the statistician's needs.

Pitfalls in the Application of Statistical Tools in Business Settings

✱Wendy Rotz, Ernst & Young LLP, 1101 NY Avenue, Washington, DC 20005, *wendy.rotz@ey.com*

Key Words: Business, Consulting, Sampling, Estimation, Quality, Education

From estimating taxes, through auditing financial statements, to quality and internal controls, businesses are increasingly applying statistical sampling and estimation methodologies. However, the statistical consultant should beware of a host of theoretical and fundamental statistical missteps common to laypersons adapting sampling methods, even when statisticians lay out the best of plans for them. Even statistical consultants are prone to occasional gaffs in the area of sampling and estimation in business settings, especially when they are either novices or working outside their field of expertise and technical training. This paper, of general interest to statistical consultants, samplers and educators, covers pitfalls identified through 17 years of experience in consulting, training, and reviewing statistical methods in the area of sampling and estimation.

Business Survey Challenges

✱Joe Callender, Ernst & Young LLP, 20005, *joe.callender@ey.com*; Amy Luo, Ernst & Young LLP

Key Words: Response Rate, Business Surveys, Benchmark Report

This paper will focus on Ernst & Young business surveys. In particular, we do many surveys whose target audience is high level executives at large companies. Conducting surveys in this environment presents unique challenges compared to household surveys or even large scale business surveys. For example, our business surveys do not have the legal mandates enjoyed by the U.S. Census Bureau and some other government agencies. This paper will review types of surveys we conduct, unique private sector challenges we face, and quality improvement efforts we have undertaken. We will discuss response issues such as low cooperation rates from hard-to-reach C-suite respondents, use of benchmark reports as incentives, and oversaturated survey environments.

Data Visualization and Analysis: Business Consulting Applications

✱Ru Sun, Ernst & Young LLP, 1101 New York Avenue, N.W., Washington, DC 20005, *Ru.Sun@ey.com*

Key Words: Visualization, Consulting, SPLUS

The QUEST (Quantitative Economics and Statistics) practice of Ernst & Young offers a full range of analytic services - from data collection to analysis to action. Our service lines include: 1) Bank Regulatory Compliance, 2) Fair Lending Regulatory/Litigation Support, 3) Federal Policy, 4) Regional Economic Contributions, 5) Risk Management & Quantification, 6) Statistical Sampling, 7) State and Local Tax Policy and 8) Surveys. This presentation will focus on how data visualization provides additional insights and understanding to the clients. Each custom-designed graph addresses a specific need of the projects/clients, whether it is a bar-plot or a map or in a new form (e.g. an animated graph). By showing patterns and comparisons, graphs help analysts to explore the trends and relationships in the data, as well as help clients to turn data into insights and make informed decisions.

503 Bayesian Applications ●

Section on Bayesian Statistical Science
Thursday, August 7, 10:30 a.m.–12:20 p.m.

Additive Models with Interval-Censored Data

✱Chin-I Cheng, University of Missouri, Department of Statistics, University of Missouri, Columbia, MO 65211-6100, *cc4f7@mizzou.edu*; Paul Speckman, University of Missouri

Key Words: additive model, interval-censored data, semiparametric

Motivated by an industrial process, this study models the measurement of wear of a die in a machine producing stamped metal parts. Regular measurements are taken on the metal parts. Because of rounding, the measurements are interval censored. Besides potential wear, other factors including time of day and operator effect are potentially present. The process is also possible cofactored including operator effect. The paper shows the results of simulation study examining the effect of interval censoring has on inference for the components of the model.

Bayesian Semiparametric Modeling of Spatially Correlated Breast Cancer Survival Incidence with Cure Fractions

✱Luyan Dai, University of Missouri-Columbia, 306 Hitt Street, Apt 3J, Columbia, MO 65201, *ld9n9@mizzou.edu*; Dongchu Sun, University of Missouri; Zhuoqiong He, University of Missouri-Columbia; Mario Schootman, Washington University

Key Words: Thin-plate Splines, Cure Models, MCMC, Survival Analysis

Cure rate models are constructed specifically for modeling time-to-event data incorporating cure fractions. We focus on the extension of cure rate models by incorporating spatially correlated frailties. In order to loose the boundary assumptions of spatial sites in commonly adopted conditional autoregressive processes (CAR), we propose nonparametric thin-plate splines for spatial effects at both census tract and county levels. The data set analyzed were recorded from year 1991 to 2000 in SEERs program in state Iowa. Different parameterizations of the smoothing parameter of the thin-plate spline prior are considered and posterior quantities are estimated by Gibbs sampling. Dimension reduction issues are also discussed to speed up the computation.

Bayes Factor Consistency for Large Model Dimensions

✱Ruixin Guo, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211, *rgd27@mizzou.edu*; Paul Speckman, University of Missouri

Key Words: Model selection, Bayes factor, Consistency, Laplace approximation

Liang et al. (2007) studied consistency of Bayes factors for model selection in linear models using proper mixtures of g-priors, where they put a flat prior for the common parameters. Marin and Robert suggest to put g-priors for both compared models so that an improper g-prior can be used in Bayes factor. In this model setup, we prove consistency of the Bayes factor under an improper g-prior (Marin-Celeux-Robert) when model dimensions are fixed. We also discuss consistency and inconsistency problems under improper and proper priors in Bayes factors when the reduced model is arbitrary with fixed dimension and the full model dimension p grows with the number of observations n. We obtain consistency and inconsistency depending on the limiting behavior of p/n.

One-Way ANOVA: Fixed Effects, or Random?

✱Fei Liu, University of Missouri, 146 Middlebush Hall, Columbia, MO 65211, *liufei@missouri.edu*; Dongchu Sun, University of Missouri; Paul Speckman, University of Missouri; Jeffrey Rouder, Univeristy of Missouri

Key Words: One-way ANOVA, Consistency, Model selection, Zellner-Siow's prior, Effective sample size

We consider two versions of the one-way ANOVA model that exist in literature, namely the fixed effect model and the random effect model. The main difference of these two types of models lies in whether or not the effects of the factor levels are treated as random variables. In practice, however, from the Bayesian point of view, all parameters are considered as random variables, making the distinction between fixed effect model and random effect model rather obscure. The primary goal of this article is to seek for a unified Bayesian approach to deal with one-way ANOVA models with fixed effects and random effects. We propose a modification of the Zellner-Siow prior, and show that the proposed prior will result good consistency properties in terms of model selection and posterior distributions, in the settings of either fixed effect model or random effect model.

Posterior Simulation in the Generalized Linear Mixed Model with Semiparametric Random Effects

✱Subharup Guha, University of Missouri-Columbia, *guhasu@missouri.edu*

Key Words: Laplace approximation, MCMC, Dirichlet process models, non-conjugate models, Polya urn scheme, semiparametric Bayesian methods

Generalized linear mixed models with semiparametric random effects are useful in a wide variety of Bayesian applications. The algorithms proposed by MacEachern and Muller (1998) and Neal (2000) are often applied to generate posterior samples in non-conjugate mixture of Dirichlet process (MDP) models. Common problems associated with these algorithms include convergence and mixing difficulties. We propose a new algorithm for MDP models with exponential family likelihoods and normal base measures. For the special case of conjugate MDP models, the algorithm is identical to the Gibbs sampler based on the Polya urn scheme. The performance of the technique is investigated using a Poisson regression model with semiparametric random effects. The technique is associated with substantial benefits relative to existing methods, both in terms of convergence properties and computational cost.

504 Applications of Nonparametric Statistics on Manifolds ●▲

Section on Nonparametric Statistics, IMS

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Nonparametric Inference on Shape Spaces

✱Abhishek Bhattacharya, The University of Arizona, Department of Math, 617 N Santa Rita Ave., P.O. Box 210089, Tucson, AZ 857210089, *abhishek@math.arizona.edu*

Key Words: shape spaces, Riemannian manifold, extrinsic and intrinsic means and variations, two-sample nonparametric tests

The statistical analysis of shape distributions based on random samples is important in many areas. To measure the shape of an object, one may pick a suitable ordered set of points on an image of the object under consideration. The equivalence class of that set of points identified modulo size, translation and rotation is called its similarity shape. In case we consider the equivalence class modulo all affine transformations, we get the affine shape of that configuration. Another notion of shape is the projective shape which is particularly appropriate in machine vision. All these shape spaces can be made into Riemannian manifolds. In this talk, I present certain recent methodologies and some new results for the statistical analysis of probability distributions on manifolds and apply them to the shape spaces to estimate shape parameters and compare different shape distributions.

Regularized Deconvolution on the Euclidean Motion Group

✱Maia Lesosky, University of Guelph, *mlesosky@uoguelph.ca*

This work solves a deconvolution problem on the Euclidean motion group. Regularization techniques are used to handle the ill-posed nature of the problem. Upper bounds for the mean integrated squared error are calculated for a number of cases. An application is presented.

Comparing Random Variables on Manifolds

✱Nikolay H. Balov, Florida State University, 117 N. Woodward Ave., 214 OSB, Tallahassee, FL 32306-4330, *balov@stat.fsu.edu*

Key Words: covariance, manifolds, comparing distributions, diffusion tensor imaging, directional data, similarity invariants

We are interested in comparing probability distributions defined on Riemannian manifolds. The traditional approach to study a distribution relies in locating a mean and finding the dispersion about the mean. On a general manifold however, even if two distributions are sufficiently concentrated and have unique means, a comparison of their covariances is not possible due to the difference in local parametrizations. To circumvent the problem we associate a covariance field with each distribution and compare them at common points by applying a similarity invariant function on their representing matrices. In this way we are able to define a distance between distributions. We discuss different choices of invariants and illustrate their properties through simulations. We also provide some results for studying Orientation Distribution Functions defined on unit sphere as appear in DTI analysis.

Applications of Nonparametric Statistics on Shape Manifolds and on Shape-and-Size Shape Manifolds

✱Ananda Bandulasiri, Sam Houston State University, Department of Mathematics & Statistics, Huntsville, TX 77340, *abandulasiri@hotmail.com*

Key Words: Shape Analysis, Extrinsic Mean, Mean Shape

Recently developed techniques in analyzing data on a 3D similarity shape manifold or on a planar shape-and-size manifold have applications in biology, medical imaging and bioinformatics. In this paper, we will discuss some nonparametric tests for extrinsic 3D mean shape for the problem of characterizing protein's active sites which are useful to understanding its biological function. In addition to this we discuss applications of 3D extrinsic shape analysis and extrinsic shape-and-size analysis in medical imaging. This is joint work with Vic Patrangenaru and Jinfeng Zhang from Florida State University.

505 Modeling Environmental Pollutants

Section on Statistics and the Environment, WNAR
Thursday, August 7, 10:30 a.m.–12:20 p.m.

Can Blood Lead Levels in Children Be Reduced?

✱Steven G. Somers, North Carolina State University, 3900A Marcom Street, Raleigh, NC 27606, *sgsomers@ncsu.edu*; Jessica P. Williams, North Carolina State University; Amanda Campbell, North Carolina State University

Key Words: children, lead, CDC, US EPA, elevated blood lead levels, statistical analysis

Lead concentration is a serious problem as there are many health hazards associated with it. Exposure in children has an adverse affect on cognitive development. Even low levels of lead intake cause damaging effects after prolonged exposure. Our objective is to investigate Estimated Blood Lead Level (EBLL) rates per 1000 children. We expect the ‘03 California EBLL data to be incorrect and will rerun previous EBLL data sets with the ‘06 data. In doing this, we hope to predict EBLL as a function of environmental factors, provide a more accurate data base for the CDC, and better inform the USEPA of areas with high EBLLs. We will investigate other states such as NC, TX, AZ, etc. to determine the integrity of their data. Exploratory statistical methods will help produce a more accurate model for EBLLs and help the government determine optimum control strategies to reduce EBLLs in children.

Can Meteorologically Adjusted Ozone Air Quality Trends Identify the Impact of the Nitrogen Oxides Utility Reductions?

✱Kristen L. Gore, North Carolina State University, Raleigh, NC 27607, *klgore@ncsu.edu*; Adrienne Wootten, North Carolina State University; Timothy Brown, North Carolina State University; Jie Zheng, North Carolina State University

Key Words: NOx, ozone, VOCs, meteorology, troposphere, NOx SIP Call

Inhalation of ground-level (tropospheric) ozone can trigger a variety of health ailments, including chest pain, throat irritation, bronchitis, emphysema, and asthma. The major precursors to ground-level ozone formation are volatile organic compounds (VOCs) and nitrogen oxides (NOx). The ability to determine the impact of ozone precursor emission controls on ground-level ozone trends is complicated by the presence of meteorological and biogenic factors. The EPA Nitrogen Oxides State Implementation Plan Call (NOx SIP Call) was implemented in 2001 in an effort to mitigate the formation of ground-level ozone. Was it successful? The purpose of this project was to build a regressive time series model that removes the effects of meteorology, autocorrelation, and seasonal trends using data from the NC Department of the Environment and Natural Resources and the NC State Climate Office.

Examining Crustal Matter: Resolving the Particulate Matter Emission Inventory/Air Quality Discrepancy

William J. Rice, North Carolina State University; ✱Stacy A. Jones, North Carolina State University, 6 Bagpipe Lane, Durham, NC 27712, *sajones4@ncsu.edu*; Lauren A. Klein, North Carolina State University

Key Words: fine particulate matter, National Emissions Inventory, PM2.5, crustal matter, ambient air, environmental factors

The crustal matter found in fine particulate matter comes from the earth’s crust, fugitive dust, which consists of unpaved roads, agricultural tilling, construction, etc. Crustal matter is affected by environmental factors such as season, wind speed, and day of the week, etc. In 2001, the National Emissions Inventory estimated that the National total percentage of crustal matter to fine particulate matter (PM2.5) was estimated at 31.4%. The ratio of crustal matter in ambient air ranges from 5-10%. This project focuses on the discrepancy in ambient air vs. emission inventory in the ratio of Crustal matter to total fine particulate matter (PM2.5). The ambient air measurements are taken from ambient air quality monitoring sites, located in both urban areas and in national parks and wilderness areas. We will discuss the model we developed to determine which variables affect this ratio.

Did the Addition of Supplementary Control Systems at Utilities Result in Lower Nitrogen Oxides Emissions and Reductions in Ground-Level Ozone in North Carolina?

✱Timothy Brown, North Carolina State University, Raleigh, NC , *tdbrown2@ncsu.edu*; Jie Zheng, North Carolina State University

Key Words: Nitrogen oxides, utilities, ozone, volatile organic compounds, meteorology, air monitoring

Nitrogen Oxides (NOx) emissions are an important pollutant that comes from combustion processes especially from transportation sources and utilities. NOx is a precursor emission, which along with volatile organic compounds (VOCs), in the presence of sunlight and warm temperatures produces ozone, the principal constituent in photochemical smog. This paper will focus on the analysis of nitrogen oxide emissions in North Carolina and Tennessee to determine the impact of reductions at utilities. Ambient ozone levels collected at four remote air monitoring sites will be analyzed to determine the impact of the NOx reductions at utilities. We will develop the model to find the trend in ozone levels in NC after the effects of meteorologically variables have been removed.

506 Surveys and Administrative Records: Data Comparison and Quality ●▲

Section on Survey Research Methods, Section on Government Statistics
Thursday, August 7, 10:30 a.m.–12:20 p.m.

Evaluating the Relationship Between Survey Design and False-Negative Reporting About Medicaid Enrollment

✱Victoria Lynch, U.S. Census Bureau, *victoria.a.lynch@census.gov*

Key Words: Record check, Medicaid, administrative records, NHIS, CPS

Record check studies show that survey respondents often provide false-negative reports to questions about Medicaid enrollment. In this study we evaluate how survey design may interact with characteristics of the enrollee to affect the probability of false-negative reporting about Medicaid in 2001. We use the same set of variables to separately model misreport using the National Health Interview Survey (NHIS) and the Current Population Survey (CPS). The results suggest that false-negative reporting is broadly similar across the surveys but survey design matters. The surveys appear to do equally well soliciting a correct report about enrollees whose period of coverage includes the date of the survey. The surveys appear to differ in how they do with enrollees from different states as well as those with private insurance or Supplemental Security Income (SSI) in the month of the survey.

Reconciling Employment Differences Between Administrative and Survey Data

Applebaum Margaret, Bureau of Labor Statistics; Kristin Fairman, Bureau of Labor Statistics; Jeffrey Groen, Bureau of Labor Statistics; ✱Polly Phipps, Bureau of Labor Statistics, 2 Massachusetts Ave NE, Room 1950, Washington, DC 20212, *phipps.polly@bls.gov*

Key Words: measurement error, survey and administrative data, employment, response analysis survey

This paper investigates differences in employment figures gathered from unemployment insurance tax filings under the Quarterly Census of Employment and Wages and employment reported through the Current Employment Statistics survey. Since these two Bureau of Labor Statistics programs both collect monthly employment from an establishment for the same reference period, the employment figures should generally be identical. However, differences exist at the micro and aggregate levels, both at a point in time and in seasonal patterns. We analyze employment differences for 200 establishments with large employment differences in 2005-06 and determine the scope of and patterns in the differences. We then report on findings from a response analysis survey in which the 200 establishments were asked about reasons for differences in employment reported on the tax and survey forms.

Comparison of Business Revenues from Two Administrative Files

✱Guylaine Dubreuil, Statistics Canada, 100, promenade Tunney's Pasture, 16-B RHC building, Ottawa, International J8Z 2N7 Canada, *guylaine.dubreuil@statcan.ca*; François Brisebois, Statistics Canada; Patrice Martineau, Statistics Canada; Julie Girard, Statistics Canada; Caroline Rondeau, Statistics Canada

Key Words: Administrative data, comparisons of variables, influential data, regression, calendarization, survey replacement

A portion of the administrative data used at Statistics Canada (STC) comes from the T2 database, which is related to corporations and is made up of financial and fiscal data.† These administrative data are provided by the Canada Revenue Agency (CRA). STC also receives from CRA information related to the Goods and Services Tax (GST) that businesses have been remitting to the CRA since 1991. The GST information provided includes total revenue and the taxes collected on products and services over a given reporting period. Most corporations are present on both sources of data and it is therefore possible to compare their revenue variables and assess the strength of the relationships between them. The presentation will focus on the different results obtained from the comparisons and discuss improvements that can be brought into the methodologies used to process these data.

Reconciling Differences in Income Estimates Derived from Survey and Tax Return Data

Kevin Moore, Federal Reserve Board of Governors; ✱Barry Johnson, Statistics of Income, IRS, P.O. Box 2608, attn: RAS:S:SS:S, Washington, DC 20013-2608, *Barry.W.Johnson@irs.gov*

Key Words: Survey data, Administrative Data, non-sampling error, Nonresponse bias

In the US, survey and administrative data sources are frequently blended to support a variety of research purposes. Because these data sources are primarily designed for different purposes, one inherently research oriented and the other to administer government programs, blending them poses unique challenges. This paper will focus specifically on income data derived from two sources: the Surveys of Consumer Finances (SCF) sponsored by the Board of Governors of the Federal Reserve System, and Federal income tax return data collected by the Statistics of Income (SOI) Division of the Internal Revenue Service. Utilizing multiple years of data, we will examine key similarities and differences between these two data sources and demonstrate methods for reconciling estimates produced from them.

507 How To Publish Your Book with the ASA-SIAM Series on Statistics and Applied Probability

Section on Statistical Education

Thursday, August 7, 10:30 a.m.–12:20 p.m.

How To Publish Your Book with the ASA-SIAM Series on Statistics and Applied Probability

✱Lisa M. LaVange, The University of North Carolina at Chapel Hill, Department of Biostatistics, CB#8030, Chapel Hill, NC 27514-3628, *lisa_lavange@unc.edu*; ✱Keith Crank, The American Statistical Association, 732 North Washington Street, Alexandria, VA 22314-1943, *keith@amstat.org*; ✱Sara J. Murphy, Society for Industrial and Applied Mathematics, 3600 University City Science Center, Philadelphia, PA 19104, *murphy@siam.org*; ✱Karen A.F. Copeland, Boulder Statistics, 2456 Vine Place, Boulder, CO 80304-0992, *karen@boulderstats.com*

This panel addresses why and how to publish your book with the ASA-SIAM Series on Statistics and Applied Probability, a publishing partnership of these two premiere societies. Issues covered will be what topics are appropriate for the series, how to submit a proposal for review, what benefits you get from publishing with the series, how the production process works, and how series books are marketed to reach the widest possible audience. In addition to the series editor-in-chief and acquisitions editor, the panel includes authors who have published in the series.

508 Causal Inference and Factor Analysis in the Social Sciences ●

Social Statistics Section, Section on Government Statistics

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Estimating Non-Numeric Outcomes with Structural Equation Modeling: Application to Colorectal Cancer

✱Ralph M. Turner, University of the Sciences, 600 S. 43rd Street, Philadelphia, PA 19104, *r.turner@usp.edu*

Key Words: Structural Equation Modeling, Logistic Regression, MCMC Estimation, Cancer Survival, Non Numeric Outcomes

Previously Structural Equation Modeling (SEM) programs required numeric scores for endogenous latent constructs making some important categorical social/health problems unsuitable for analysis. However, simpler access to Markov chain Monte Carlo (MCMC) simulation techniques make it possible to draw random values of parameters for logistic weights and survival time, from high-dimensional joint posterior distributions. The present study investigated the long-term functional status of 177 colorectal cancer patients following colostomy using SEM augmented with MCMC Bayesian estimation. A structural model designed to simultaneously assess the comparative importance of Optimism, Expectancies, Depression, and Symptom Distress on Functional Dependency at three and six-month post-treatment illustrates the utility of this development. Optimistic expectations and mood most influence recovery.

A Parametric Mixture Model Method for Clustering Multivariate Correlated Binary Data

✱Ajit C. Tamhane, Northwestern University, Department of IEMS, 2145 Sheridan Road, Evanston, IL 60208-3119, *atamhane@northwestern.edu*; Dingxi Qiu, University of Miami

Key Words: Latent class analysis, Clustering, Multivariate Bernoulli Data, Maximum likelihood, Bayes classification rule, Mixture model

The traditional latent class analysis (LCA) uses a binary mixture model with independent responses. In many practical applications, the responses are locally dependent because they are observed on the same subject. We extend the LCA model to allow for local dependence by fitting a parametric mixture model in which each cluster follows a multivariate Bernoulli distribution. An extension of a family of parametric models by Oman and Zucker (2001) is adopted for this purpose and the method of maximum likelihood estimation is used for fitting. The Bayesian information criterion (BIC) due to Schwarz (1978) is employed to select the number of clusters. Subjects are classified to clusters using the Bayes rule. The proposed mixture model method is illustrated by applying it to two real data sets.

Recommended Sample Size for Conducting Exploratory Factor Analysis on Dichotomous Data

✱Robert Pearson, University of Northern Colorado, Campus Box 124, University of Northern Colorado, Greeley, CO 80639, *pear7541@unco.edu*; Daniel Mundfrom, University of Northern Colorado

Key Words: Factor Analysis, Sample Size, Multivariate, Simulation, Categorical

Exploratory factor analysis is used to identify and interpret correlational structures that may be present in multivariate data. One effect of sampling error is a decrease in agreement between estimated factor loadings and population values. The purpose of this study was to identify the minimum necessary sample size needed to adequately reproduce population loadings when binary data are analyzed. A Monte Carlo simulation was conducted, varying the level of communalities, number of factors, variable-to-factor ratio, and dichotomization threshold. The degree of agreement between population and sample loadings was measured using a coefficient of congruence and interpreted using criteria suggested by Tucker, Koopman, and Linn (1969). Results mimic patterns observed for continuous data by Mundfrom et al. (2005), but with higher necessary sample sizes and more extreme trends.

The Causal Effect of Class Size on Academic Performance: Multivariate Instrumental Variable Estimators with Tennessee Class Size Data Missing at Random

✱Yongyun Shin, The University of Chicago, Department of Sociology, Room 417, 1126 E. 59th Street, Chicago, IL 60637, *yongyun@uchicago.edu*; Stephen W. Raudenbush, The University of Chicago

Key Words: Causal Effect, Class Size, Ignorably Missing, Instrumental Variable, Maximum Likelihood, Simultaneous Equation Model

Does reduced class size cause higher academic achievement in reading, mathematics, listening and word recognition skills? How large are the effects? Do they vary across schools? To address the questions, we analyze data from Tennessee's Student/Teacher Achievement Ratio study of 1985. We estimate a three-level multivariate simultaneous equation model with an instrumental variable (IV) via maximum likelihood to analyze the data. The IV, random assignment of students to a small or regular class, reduces class size which by hypothesis improves academic achievement. We extend single-level Rubin's Causal Model to three levels. The method handles ignorably missing data with a general missing pattern using all available data. The results show that reduced class size improves the four test scores from kindergarten to third grade. We find that the causal effects of class size are homogeneous.

Three-Mode Models for Multitrait-Multimethod Data

✱Frans J. Oort, University of Amsterdam, Department of Education, Nieuwe Prinsengracht 130, Amsterdam, 1018VZ The Netherlands, *f.j.oort@uva.nl*

Key Words: factor analysis, structural equation modeling, covariance structure analysis, multitrait-multimethod

Multitrait-multimethod (MTMM) data are characterized by three modes: traits, methods, and subjects. Considering subjects as random, and traits and methods as fixed, stochastic three-mode (S3M) models can be used to analyze MTMM covariance data. S3M models can be written as linear latent variable models with direct product restrictions on the parameter matrices (Oort, 1999), yielding three-mode factor models (Bentler & Lee, 1979) and composite direct product models (Browne, 1984) as special cases. Direct product restrictions on factor loadings and factor correlations facilitate interpretation of the results and enable easy evaluation of the validity requirements of MTMM correlations (Campbell & Fiske, 1959). As an illustrative example, a series of S3M models has been fitted to data of three personality traits of 482 students, measured with twelve items, through three methods.

509 Clinical Trials Methodology

Biometrics Section, Section on Nonparametric Statistics, Biopharmaceutical Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Estimation Method of the Semiparametric Mixture Cure Gamma Frailty Model

✱Jiajia Zhang, University of South Carolina, 800 sumter street, Columbia, SC 29208, *jzhang@gwm.sc.edu*; Yingwei Peng, Queen's University

Key Words: Mixture Cure Model, Frailty Model, EM algorithm, Multiple Imputation

Mixture cure frailty model has been proposed to analyze censored survival data with a cured fraction and unobservable information among the uncured patients. Different from a usual mixture cure model, the frailty model is employed to model the latency component in the mixture cure frailty model. In this paper, we extend the mixture cure frailty model by incorporating covariates into both the cure rate and the latency distribution parts of the model, and propose a semiparametric estimation method for the model. The EM algorithm and the multiple imputation method are employed to estimate parameters of interest. In the simulation study, we show that both estimation methods work well. To illustrate, we apply the model and the proposed methods to a data set of failure times from bone marrow transplant patients.

Adaptive Clinical Trial Designs To Meet the Criteria

✱Yi Cheng, Indiana University South Bend, 1700 Mishawaka Ave., South Bend, IN 46634, *ycheng@iusb.edu*

Key Words: Bayesian method, Clinical Trials, Group Sequential designs, Adaptive designs, Utility function, Randomization

Optimal decision-analytic designs are deterministic. Such designs are appropriately criticized in the context of clinical trials because they are subject to assignment bias. On the other hand, balanced randomized designs may assign an excessive number of patients to a treatment arm that is performing relatively poorly A compromise between these two extremes, one that achieves some of the good characteristics of both, is desirable, which leads to the designs to be proposed in this talk. The extensions to various clinical trial settings will also be discussed.

Using Short-Term Response To Facilitate Adaptive Randomization for Survival Clinical Trials

✱Xuelin Huang, The University of Texas M.D. Anderson Cancer Center, 1515 holcombe BLVD, Unit 447, Houston, TX 77030, *xlhuang@mdanderson.org*; Jing Ning, The University of Texas M.D. Anderson Cancer Center; Yisheng Li, The University of Texas M.D. Anderson Cancer Center; Donald A. Berry, M.D. Anderson Cancer Center

Key Words: Adaptive randomization, Bayesian, Clinical trial design, Interim monitoring, Survival analysis

It has always been difficult to design and implement a response-adaptive randomization for survival clinical trials due to the long waiting time to observe the survival outcome. In practice, often information about a short-term response is quickly available during or shortly after treatment, and this short-term response is a good predictor for long-term survival. In this article, we propose a new design for survival trials which can utilize short-term response information to “speed-up” the response-adaptiveness of the randomization procedure. A connection between short-term response and long-term survival is established, first by prior clinical information, and then updated dynamically by the information accumulated in the current trial. Interim monitoring and final decision are based upon inference on the primary outcome, namely, survival.

Nonparametric Adaptive Sample Size Calculation in Diagnostic Trials

✱Liansheng Tang, George Mason University, 22030, *ltang1@gmu.edu*

Key Words: ROC, Adaptive trial, diagnostic medicine

In diagnostic trials, required sample sizes have to been calculated based on the variance of the statistic of interest, which is usually the difference between the area or partial area under the ROC curve for the two diagnostic methods. Without pilot data in comparative diagnostic trials, the variance of a statistic is specified by assuming parametric models for nondiseased and diseased subjects. However, if parametric models are not correctly specified, the required sample size might be smaller than necessary and it will result in an underpowered study. In this paper we propose a nonparametric adaptive method to adjust sample sizes during diagnostic trials. Our method is shown to maintain the power and type I error. The finite sample property of our method is evaluated through large-scale simulation studies.

Sample Size Re-Estimation in an Adjuvant Cancer Trial

✱Erinn M. Hade, The Ohio State Univeristy, M200 Starling Loving Hall, 320 W 10th Ave, Columbus, OH 43210, *erinn.hade@osumc.edu*; David Jarjoura, The Ohio State Univeristy

Key Words: sample size, re-estimation, cancer, re-sampling

Various methods for re-evaluation of a randomized trial’s sample size have been proposed, based on the recognition that the information used to design the study can be updated to avoid inadequate power. With long accrual times typical of a cancer survival trial, data are available for re-estimating failure rates. Although sample size estimates depend mainly on the proposed effect of treatment on the difference in failure rates, a mistakenly high estimate of the failure rate at the design stage may reduce the power unacceptably for a clinically important hazard ratio. In this presentation we propose a method to avoid this by re-estimating the failure rate.

510 Model-Based Estimation

Section on Survey Research Methods, Section on Nonparametric Statistics, Section on Government Statistics, Social Statistics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Inference in Finite Population Sampling Using Regression Models Imposed by Randomization in the Sample Design

✱Steve Woodruff, Specified Designs, 800 West View Terrace, Alexandria, VA 22301-2750, *stevewoodruff@comcast.net*

Key Words: Probability Sampling, Regression Models, Stochastic Process, GLS & BLUE Estimation

The variance of a Horvitz-Thompson based estimate is sensitive to the sample design and is often quite large for at least some of the study variable estimates. Models and model based estimates can provide a more efficient alternative but are dependent on possibly questionable model assumptions and can sacrifice the impartiality of randomization. For some sampling problems there is a third estimation technique which is implemented by appending randomized construction of sample units to the usual design for selecting sample units. This technique imposes a model, retains the comforting impartiality of randomization and the optimality of a Best Linear Unbiased Estimator (BLUE) under a model that is assured by this additional design feature. Theory and examples are presented for complex stratified cluster designs where reduction in sampling variance ranges to over a hundredfold.

Robust Predictions Based on Model-Based Approaches

✱Bing Han, RAND Corporation, 90401, *bhan@rand.org*

Key Words: Residual, Model-based approach, Robust prediction

We discuss a robust prediction tool based on the residuals of a model-based approach to survey data. The observed residuals can be used to estimate the error distribution under the i.i.d. assumption. The resulted prediction interval has better accuracy than the asymptotic normal or t prediction interval. We illustrate the potential by the analysis of historical trend of obesity in American young adult from the national health survey BRFSS.

A Snowball’s Chance in Nigeria

✱Rachel Schutt, Columbia University, Statistics Department, *rrs2117@columbia.edu*; Alexandra Scacco, Columbia University

Key Words: Snowball Sampling, Respondent Driven Sampling, Bayesian, Network, Regression

Motivated by a desire to better understand the riots that took place in towns in Nigeria in 2000 and 2001, two simultaneous surveys were undertaken in Nigeria in the Fall of 2007. One set of survey data was collected using a random sample, and the other was collected using a variant of snowball sampling, Respondent Driven Sampling (RDS). Using both sampling methods allows us to leverage information from the random sample to better understand and adjust for bias in the RDS sample. Viewing the analysis of snowball sampled data as a network analysis lying on the continuum between a local and complete network analysis, we are developing a Bayesian regression-type analysis for use with snowball data, as well as exploring other methods for making statistical inferences.

The Two-Sample Problem

✱Alan H. Dorfman, Bureau of Labor Statistics, Washington, DC 20212, *dorfman.alan@bls.gov*

Key Words: Missing Information Principle, Post-stratification, Pseudo-likelihood

It sometimes happens that two separate samples from a population, having perhaps quite distinct designs and mode of sampling, contribute information on the same variable of interest, and it becomes an important question how to combine the data from the two samples. An example is the Occupational Employment Statistics survey (OES) and the National Compensation Survey (NCS), carried out by the Bureau of Labor Statistics, both contributing information on occupational wages. We discuss some new options for combining data from two samples and achieving unified estimation.

Robust Improvements in Ratio-Type Estimators

✱Evrim Oral, Middle East Technical University, Middle East Technical University, Department of Statistics, Ankara, 06531 Turkey, *eoral@metu.edu.tr*; Cem Kadilar, Hacettepe University

Key Words: Ratio-type estimators, Simple random sampling, Robust regression, Modified maximum likelihood methodology

In sampling theory, when the correlation between study and auxiliary variables is positively high, the classical ratio estimator (RE) is the most practicable estimator to estimate the population mean. Sisodio and Dwivedi (1981) and Upadhyaya and Singh (1999) suggested to use many population information of the auxiliary variable to increase the efficiency of the RE. Kadilar and Cingi (2004) combined this suggestion with the estimator in Ray and Singh (1981) and proposed novel ratio estimators. In this study, we adapt robust regression to the Kadilar-Cingi estimators (KCE) and obtain the conditions where these adapted estimators are more efficient than KCE theoretically. We support theoretical results with simulations and compare the adapted estimators with both KCE and the classical RE. We study the robustness properties of the new estimators via simulations and give a real-life example.

Improved Ratio Estimators in Adaptive Cluster Sampling

✱Feng Min Lin, National Cheng Kung University, Department of Statistics, No 1 TaHsueh Road, Tainan, 70101 Taiwan, *fmlin@stat.ncku.edu.tw*; Tzu Ching Chiang, National Cheng Kung University

Key Words: Auxiliary Variable, Ratio Estimator, Adaptive Cluster Sampling, Rao-Blackwell Theorem

For better inference of the population quantity of interest, ratio estimators are often recommended when certain auxiliary variables are available. Two ratio estimators in adaptive cluster sampling have been studied in Dryver and Chao (2007); however, they are not functions of the minimal sufficient statistics for this sampling design and hence can be improved. The object of this study is to apply Rao-Blackwell theorem to obtain more efficient and easy-to-compute estimators. The improved ratio estimators obtained by conditioning on sufficient statistics are described. Despite having higher variances, the improved estimators are easier to be computed than estimators obtained by conditioning on the minimal sufficient statistics. The performances of the improved ratio estimators are illustrated in simulation studies.

511 Analytic Methods with Noncontinuous Outcomes

Section on Survey Research Methods, Section on Nonparametric Statistics, Section on Government Statistics

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Estimating the Distribution Function Using Ranked Set Samples with Imperfect Ranking

✱Satrajit Roychoudhury, Schering-Plough Research Institute, KENILWORTH, NJ 07033, *satrajit.roychoudhury@spcorp.com*; Kaushik Ghosh, University of Nevada, Las Vegas

Key Words: Rank set Sampling, Imperfect Ranking, Density Estimation

Ranked set sampling is a technique of efficient data collection when measuring of units (on the variable of interest) is expensive or time consuming but one can easily rank them. In practical situations however, any such ranking procedure employs visual or surrogate-based ranking, which may be prone to errors, giving rise to biased inferences. In this paper, a method has been developed to estimate the underlying true sampling distribution in the presence of ranking errors that automatically corrects for the bias introduced. Asymptotic results are derived and are illustrated with a numerical study.

Power Analysis for the Regression Discontinuity Design

✱Hyunshik Lee, Westat, 1650 Research Blvd., Rockville, MD 20850, *HyunshikLee@westat.com*; Tom Munk, Westat

Key Words: Effect size, Random experimental design, Inflation factor

In many education and social intervention programs it is thought efficacious and ethical to offer the treatment to individuals in greatest need. This entails assigning subjects to the treatment group if they are at or above some cutoff score on a measure related to need. The outcome of the treatment is then analyzed using the regression discontinuity design technique assuming that the treatment effect can be detected at the assignment cutoff through regression analysis. In this study the power of such a design is analyzed to determine the sample size necessary for implementing the parametric regression model. It was found that higher order terms of the score variable and interaction terms with other covariates in the model have a damaging effect on power. Some strategies to handle these effects are discussed.

Charactering the Propensity To Volunteer in America

✱Shelton Jones, RTI International, PO Box 12194, Research Triangle Park, NC 27709, *smj@rti.org*; David Weitzenkamp, RTI International; Kathleen Jordan, RTI International; Jeniffer Iriondo-Perez, RTI International; John Foster-Bey, CNCS; Annette Green, RTI International

Key Words: synthetic design variables, pseudo design variables, volunteering

The Volunteer Supplement of the Current Population Survey database is used to answer questions relative to the volunteering of the American Baby Boomer population. This population consists of individuals born between 1946 and 1964. As of 2006, persons in this age group represents more than 26 percent of the U.S. population and obviously makes considerable contributions to many organizations and services by volunteering. Using the Volunteer Supplement, we will present many characteristics of Baby Boomer volunteering, describe distinctions between them and other generations, and describe how their volunteering is changing over time. Synthetic design variables were created to estimate sampling variance and logistic regression used to show that strong predictors of volunteerism can be determined from home ownership, business ownership, and voting frequency.

B-Splines and Bootstrapping for Piecewise Logistic Regression in Complex Samples

✱Scott W. Keith, The University of Alabama at Birmingham, Ryals Public Health Bld, 1665 University Blvd, Birmingham, AL 35294, *skeith@ms.soph.uab.edu*; David B. Allison, The University of Alabama at Birmingham

Key Words: free-knot, B-splines, bootstrap, multi-stage samples, logistic regression, nonlinear models

Splines are a flexible application for nonlinear modeling. Estimating both the number and locations of knots as free parameters (join points in a "free-knot" spline) can optimize model parsimony. We designed a spline framework that estimates nonlinear relationships between a binary outcome variable and a continuous prognostic variable in the presence of covariates in complex nationally representative samples. We use direct search methods to maximize likelihood equations with piecewise linear free-knot B-splines. Model selection and inference incorporate parametric and nonparametric bootstrap methodology. Parameter estimates are structured for interpretability so that results look similar to a logistic regression. Unlike other nonlinear approaches, our framework handles complex multistage sampled data. Our presentation will include a summary of methodological details and simulations.

Inference on Polychotomous Responses in Finite Population: A Predictive Approach

✱Tathagata Bandyopadhyay, Indian Institute of Management Ahmedabad, Wing 12, Indian Institute of Management Ahmedabad, Vastrapur, Ahmedabad, 380015 India, *tathagata@iimahd.ernet.in*; Sumanta Adhyay, University of Calcutta

Key Words: Multinomial logit model, Design-based estimator, Model-based estimator, Predictive approach

The problem of estimation of population proportions of two or more categories based on a probability sample from a finite population is considered when auxiliary information is available for all the population units. The naïve design-based estimator does not make use of this auxiliary information at the estimation stage. We adopt here a model-based predictive approach to enhance the efficiency of the estimators. We consider a multinomial logit type model where the logit function, instead of just being linear, is a known parametric function of the covariates. We use this model for the prediction of non-sample responses which are then combined with the sample responses to obtain the estimates of the finite population proportions. The expressions for asymptotic biases and variances of these estimators are obtained. An extensive data analysis is given.

512 Linear Model Extensions and Related Multivariate Problems

Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Multilevel Overdispersion in Hierarchical Generalized Linear Models

✱Trent L. Lalonde, Arizona State University, Tempe, AZ 85281, *lalonde@mathpost.asu.edu*; Jeffrey R. Wilson, Arizona State University

Key Words: Overdispersion, Generalized Linear Model, Hierarchical, Multilevel

Overdispersion is a common problem in generalized linear models with categorical response variables. We can correct for overdispersion either by including random effects in the systematic component of the model, or by adjusting the distributional assumptions to allow for greater variation. In hierarchical generalized linear models overdispersion is typically accounted for only at the response level. We suggest that overdispersion may exist at many levels of the model and should be accounted for separately at each level. This paper discusses methods for correcting for overdispersion at multiple levels by adjusting the distributional assumptions differently at each level. We include extra dispersion parameters to account for greater variation and also assume no specific distribution for the response and random effects. Consequently, estimation uses extended quasi-likelihood methods.

Stability Measures of Samples in Clustering Analysis Highly Correlates with Quality Assessment Metrics for Affymetrix GeneChip Data

✱Fenghai Duan, University of Nebraska Medical Center, 984375 Nebraska Medical Center, Omaha, NE 06410, *fduan@unmc.edu*

Key Words: Affymetrix, clustering, stability, quality assessment

In microarray experiments, unsupervised clustering algorithms are commonly implemented to reveal hidden clusters that might be relevant to disease risk structures. Methods have recently been developed to assess clustering reliability in order to derive stable clusters. On the other hand, quality assessment is an important procedure to flag outlier arrays in microarray experiments. It is of interest to study the relation of these two approaches for sake of identifying outlier arrays. To address this issue, we developed methodology to quantify the statistical significance of stability measures in clustering reliability analysis. The comparison of this method to various Affymetrix's quality assessment (QA) metrics shows that they are highly correlated regarding isolating outlier arrays. The results of this study can help specify the cutoff values for using Affymetrix's QA metrics.

Repeated Measurements on Distinct Scales with Censoring

✱Tanzy Love, Carnegie Mellon University, Dept of Statistics, Pittsburgh, PA 15213, *tanzy@andrew.cmu.edu*; Alicia Carriquiry, Iowa State University

Key Words: multiscale, cDNA arrays, gene expression, hierarchical models, preprocessing

Microarray data are subject to multiple sources of measurement error. One source of potentially significant error is the settings of the instruments (laser and sensor) that are used to obtain the measurements of gene expression. Because `optimal' settings may vary from slide to slide, operators generally scan each slide multiple times. The multiple readings are repeated measurements on different scales with differing censoring. Typically the operators then choose the reading with the fewest overexposed and underexposed spots. We propose a hierarchical modeling approach to estimating gene expression that combines all available readings on each spot and accounts for censoring in the observed values. The basic premise is that all readings contribute some information about gene expression and that after appropriate rescaling, it is possible to combine all readings into a single estimate.

Bootstrapping Methods for Data with Multiple Levels of Variation

✱Zhen Pang, Nanyang Technological University, Division of Mathematical Sciences, Singapore, International 637616, *ZPang@ntu.edu.sg*; Christopher Field, Dalhousie University; Alan Welsh, The Australian National University

Key Words: hierarchical data, mixed models, random effects, unbalanced data

We consider the problem of bootstrapping data with multiple levels of variation. We discuss the random effect model and the transformation model for such data and consider general estimators including the Gaussian maximum likelihood estimator for the mean and variance parameters of the model. We show that these estimators have different distributions under the two models unless all the random variables have Gaussian distributions. This means that, in general, inference, including bootstrap inference, depends on which model holds and we have to use procedures which are appropriate to the underlying model. We investigate the asymptotic properties of bootstrap procedures designed for the two models and discuss conditions under which the bootstrap gives asymptotically valid results.

P-Value Distribution for Linear Models in Microarray Analysis and Related Problems

✱Natalia F. Tchetcherina, The Pennsylvania State University, 514 Wartick bld, University Park, State College, PA 16802, *nxt7@psu.edu*

Key Words: Analysis of microarray data, p-values distribution, linear models

Microarray analysis is a powerful tool of genetic research. It allows researchers to test thousands of hypotheses simultaneously. Linear models are a very popular way of analysis microarray data to their simplicity and generality. Since the number of hypothesis is large Type I errors become an issue. There are several methods to address this problem most popular of which is the control of the false discovery rate. However all the methods of controlling Type I error assume that the null hypothesis distribution is explicitly known that is that the p-values are uniformly distributed. On the other hand several publications reveal a nonuniform distribution of p-values. In this paper we investigate the reasons of nonuniformity of the p-values and develop some techniques for correcting them. Our analysis based on Arabidopsis data suggests that the main reason for nonuniformity of the p-value

Differential Gene Expression Analysis: Principal Component Analysis and Biplot

✱Nusrat Jahan, James Madison University, MSC 1911, 319 Roop Hall, Department of Mathematics and Statistics, Harrisonburg, VA 22807, *jahannx@jmu.edu*; Louise Temple-Roserbrook, James Madison University

Key Words: differential gene expression, microarray, normalization, principal component analysis, biplot

A two-channel microarray experiment involving two strains of Salmonella enterica serovar Typhimurium is studied to detect genes that show differential expression across the two strains. Null hypothesis of no differential gene expression is tested for the normalized data. The differentially expressed genes are further investigated by using principal component analysis and biplots to confirm and interpret their differential status.

513 Linkage Analysis ●

Section on Statistics in Epidemiology, Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

A Note on the Asymptotic Null Distribution of Likelihood-Ratio Tests for Multivariate Genetic Linkage in Variance Components Models

✱Summer S. Han, Yale University, 24 Hillhouse Avenue, Department of Statistics, New Haven, CT 06520, *seongmin.han@yale.edu*; Joseph T. Chang, Yale University

Key Words: Likelihood-ratio test, boundary condition, mixing probability, asymptotic null distribution, multivariate genetic linkage

This study concerns the asymptotic null distribution of likelihood ratio tests for detecting genetic linkage in multivariate variance component models. In a number of previous papers, the asymptotic null distribution has been stated to be a mixture of several chi-squared distributions, with binomial mixing probabilities. We performed simulations for null distributions of k-trait likelihood ratio tests (k=2, 3, 4, 5) and found that the distribution is not actually a mixture of chi-square distributions and the mixing proportions are not binomial, giving the highest probability to the case where all variance parameters are estimated nonzero. Correcting the null distribution gives more conservative critical values than previously stated, yielding P values up to 10 times larger. Geometric arguments are given to explain the results.

Identification of Disease Loci for Schizophrenia by a Family-Based Dynamic Association Approach

✱Tun-Hsiang Yang, University of California, Los Angeles, Statistics, 11140 Rose Ave. 104, Los Angeles, CA 90034, *yangth@ucla.edu*; Kerchau Li, University of California, Los Angeles; Chih-Min Liu, National Taiwan University Hospital; Hai-Gwo Hwu, National Taiwan University Hospital

Key Words: Genome-wide linkage analyses, Schizophrenia, interactions

Genome-wide linkage analyses of complex traits have only been modestly successful in identifying regions that may harbor susceptibility disease genes. Here we developed a statistical method, Family based Dynamic Association (FDA), to augment the Haseman-Elston (HE) regression method for detecting locus interactions in nuclear family data. FDA uses the squared trait differences (Y), IBD sharing at locus 1 (X) and IBD sharing at locus 2 (Z) to assess the effect on the HE analysis of Y against X due to conditioning on Z. Our method combine the merit of allele sharing method to detect linkage at one region and take the likely interactions from the rest of genome into account. In this paper, we apply our method to the schizophrenia data. Our most significant finding on chromosome 10 at 117 cM confirmed with some recently reported regions that linked to the schizophrenia on chromosome 10.

Genetic Linkage Analysis Accounting for Missing Phenotypes in Males

✱Rosa C. Banuelos, Rice University, 2630 Bissonnet St, Apt 167, Houston, TX 77005, *banuelos@rice.edu*

Key Words: breast density, phenotype, simulation, linkage analysis

Epidemiological studies consistently report mammographic breast density as an important risk factor for breast cancer yet the etiology of increased breast density is not completely understood. Missing phenotype data in

pedigrees can potentially reduce the statistical power of linkage analysis. Usual approaches for dealing with missing data are to remove the incomplete observations or to impute missing values. However, the assumptions of these techniques of missing at random (MAR) may not always hold. In the breast density studies, the phenotype is not measured in men. Through simulation, we study the statistical properties of several missing data techniques in the genetic linkage setting.

A Variance Components Method To Test for QTL Heterogeneity Among Multiple Ethnic Groups

✱Trecia A. Kippola, Oklahoma State University, MSCS 301, Dept of Statistics, Stillwater, OK 74078, *trecia.kippola@okstate.edu*; Karen L. Edwards, University of Washington; Stephanie A. Monks, Oklahoma State University

Key Words: heterogeneity, variance components, QTL analysis, multiple ethnic groups, metabolic syndrome

Numerous methods exist for conducting linkage analysis for a quantitative trait locus (QTL); however, these methods do not allow for heterogeneity that may arise when testing for linkage using multiple families from different ethnic groups, countries or geographic locations. We propose a method, which is an extension of the variance components method suggested by Almasy and Blangero (1998). Our method can determine if there exists statistically significant linkage while allowing for heterogeneity among different ethnic groups, countries or geographic locations. We also propose a test for determining if heterogeneity exists. Both tests are based on likelihood ratio test statistics and will be applied to an existing data set consisting of genotype and metabolic syndrome phenotype data. Simulation results of various statistical properties, of our proposed methods will be presented.

Detecting and Controlling for Population Stratification in Linkage Analysis

✱Guan Xing, Bristol-Myers Squibb Company, 311 Pennington-Rocky Hill Road, Pennington, NJ 08534, *guan.xing@bms.com*; Chao Xing, The University of Texas Southwestern Medical Center

Key Words: linkage analysis, population stratification, model-free

Population stratification directly affects allele frequency estimation at the population level, and subsequently has an impact on linkage analysis depending on the study design. In this study, by taking an analogy between genetic linkage and association analysis we elucidate the role of population stratification in linkage. Focusing on model-free linkage analysis of sibship data on a binary trait we propose detecting population stratification by genotyping additional members other than the affected in a proportion of families and comparing the proportion of alleles shared identical-by-descent estimated from large families with that estimated from small families, and propose controlling for it by a matched study design with paired comparison tests.

514 Inference

Section on Statistical Computing

Thursday, August 7, 10:30 a.m.–12:20 p.m.

On Sampling and Spurious Granger Causality

✱Victor Solo, University of New South Wales, Kensington, Sydney, International 2073 Australia, *v.solo@unsw.edu.au*

Key Words: time series, granger causality

We discuss the impact of sampling and additive noise on Granger causality. It is well known that in general a subsampled vector time-series (VTS) model is a vector ARMA (VARMA) process; but there has been no general procedure for computing it. Using state space methods based on the algebraic Ricati equation we provide such an algorithm. We then give examples to show that spurious Granger causality (GC) can arise due to slow sampling. We also show spurious GC can arise due to additive noise. The consequences of these results are profound. They suggest that without prior knowledge of the time scale of causality the results of any GC analysis must be in doubt. It is also implicit that GC analysis can only be done with VARMA models not with VAR models as is commonly the case.

On Robust Estimation of the Heteroscedasticity Covariance Matrix

✱Tsung-Chi Cheng, National Cheng-Chi University, 64 ZhihNan Road, Section 2, Taipei, 11605 Taiwan, *chengt@nccu.edu.tw*

Key Words: Heteroscedasticity, L1 estimator, robust diagnostics

The assumption of homogenous variance in the normal regression model is not always appropriate. The invalidity of standard inference procedure may be produced due to the wrong estimation of the standard error when the disturbance process in a regression model presents heteroscedasticity. Test based on a heteroscedasticity consistent covariance matrix (short for HCCM) estimator is popular in application because there is no need to specify the structural form of heteroscedasticity and it is easy to compute (White 1985). The leverage points are decisive for the finite sample behavior than the degree of heteroscedasticity in the estimation of HCCM (Cribari-Neto and Zarkos 2001; and Cribari-Neto 2004). In this paper we propose a robust estimator for the heteroscedasticity covariance matrix, which is based on the concept of robust weighted L1 estimates.

Multisample Tests of Equality for Multidimensional Data with Applications to Detecting Changes in Data Streams

✱Adam Petrie, University of Tennessee, Knoxville, 3100 Lake Brook Blvd, Apt 63, Knoxville, TN 37909, *apetrie@utk.edu*

Key Words: Multisample test of equality, Spanning tree, Multidimensional test of equality, Data Streams, Detecting Change

Testing whether two or more samples appear to have a common distribution is a classic problems in statistics. Classical methods assume normality and test for the equality of means (ANOVA for d=1, MANOVA for d>1), while nonparametric tests may compare ranks when d=1 or pairwise interpoint distances if d>1. However, distributions can differ in more than just their moments. Friedman and Rafsky introduced a nonparametric graph theoretic two-sample test of equality. We generalize the asymptotic distribution of the test statistic for an arbitrary number of samples, and compare its power against other nonparametric tests based on nearest neighbor and energy statistics. Our test is more powerful when samples differ in size, location, and scale simultaneously. Further, we discuss applications of our multisample test for detecting general distributional changes in multidimensional data streams.

A Fast Algorithm for the Nonparametric MLE of the Distribution Function with Censored Data

✱Yaming Yu, University of California, Irvine, CA 92697, *yamingy@uci.edu*

Key Words: bivariate censored data, doubly censored data, EM, ICM, vertex exchange method

A fast and easy-to-implement algorithm is proposed to compute the nonparametric maximum likelihood estimator (NPMLE) for the distribution function with censored data. Based on a simple enhancement of expectation-maximization (EM) and the vertex exchange method (VEM), this algorithm is competitive with the very effective EM/ICM algorithm of Wellner

and Zhan (1997) in the univariate case, and far superior to either EM or VEM in the case of bivariate interval censored data. The method is illustrated by simulations and real data from an AIDS study.

On Model-Based Approaches for Detecting Differentially Expressed Genes and Estimating the FDR in Replicated Microarray Data

✱Shunpu Zhang, University of Nebraska Lincoln, Department of Statistics, 340 Hardin Hall North, Lincoln, NE 68583-0963, *szhang3@unl.edu*

Key Words: microarray, differentially expressed genes, mixture model method, t-mixture model, false discovery rate

Genome-wide expression data generated from the microarray experiments are widely used to uncover the functional roles of different genes. A key step to achieve this is to identify the differentially expressed genes. In this paper, we focus on a recently proposed nonparametric approach—the mixture model method. We show that the existing mixture model method does not fully use the potentials of the model based approach and has the same problems as the non-model based empirical methods. In this article, we propose a new model based method for both detecting DE genes and estimating the false discovery rate (FDR). The new method is based on the t - mixture model and provides continuous FDR estimates. Finally, the proposed method will be evaluated using extensive simulations and real microarray data.

Improved Confidence Regions for Seemingly Unrelated Regression Models

✱Kent Riggs, Stephen F. Austin State University, Department Of Mathematics and Statistics, Box 13040 SFA Station, Nacogdoches, TX 75962-3040, *riggske@sfasu.edu*

Key Words: Seemingly Unrelated Regression, Variance-Covariance Reduction, Bartlett-type correction

We employ maximum likelihood (ML) methods under the seemingly unrelated regression (SUR) model to derive five new confidence ellipsoids for covariate coefficient vectors of interest. The new confidence ellipsoids include a Bartlett-type corrected percentile on a Wald statistic and a Wald statistic with the confidence coefficient determined by a parametric bootstrap. Lastly, we study the behavior of our new confidence ellipsoids and the ordinary least squares (OLS) ellipsoid in a Monte Carlo simulation that demonstrates the improvement of several new confidence ellipsoids over the OLS method. For the configurations studied in our simulation, we determine that the Wald statistic with a Bartlett-corrected confidence coefficient is preferred in that it has close to nominal coverage and relatively small volume. Finally, we apply two of the new confidence ellipsoids to a real data set.

Sample Size Estimation Tables for Effect Size Values for the Significant Interaction Effect of a Completely Randomized Factorial Design (CRF-pq): Fixed Effects and Random Effects

✱Larry W. Featherston, University of Arkansas - NORMES, 346 West Avenue, #302, Fayetteville, AR 72701, *lfeathe@uark.edu*; Calli Johnson, University of Arkansas - NORMES; Charles Stegman, University of Arkansas - NORMES; Denise T. Airola, University of Arkansas - NORMES

Key Words: Power, Effect Size, Sample size, CRF-pq, Completely Randomized Factorial Designs, Randomized Block Designs

The purpose of this paper is to extend the work of Kirk (1995) and Foster's (1993) on estimating sample sizes for effect size values of Completely Randomized and Randomized Block designs. While Foster's sample size estimation tables are an invaluable resource for researchers using simple research designs, no such resource exist for researchers who are using more complex research designs. Therefore, this paper provides an explanation of the calculations and sample size estimation tables for the signification interaction effects of a Completely Randomized Factorial (CRF-pq) designs for both Fixed and Random Effects Models. Sample size estimation tables have been completed for CRF-22 through CRF-35 designs.

515 Inspection, Capability, and Classification Analysis

Section on Quality and Productivity, Section on Physical and Engineering Sciences

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Assessing a Manufacturing Customer's Complaint Using Supplier Process Data

✱Jon M. Lindenauer, Weyerhaeuser, 32910 Weyerhaeuser Way S, Federal Way, WA 98001, *jon.lindenauer@weyerhaeuser.com*

Key Words: Partial Least Squares-Discriminant Analysis, Multivariate Analysis, Score Plot, Loading Plot, Variable Importance Plot, Simca-P

A manufacturing plant has had periodic run rate issues with a supplier's product. From time to time, the supplier product would cause the customer to slow down their machines. The customer identified supplier product deemed as having "good" or "bad" runability. The process data associated with these runs was extracted from the supplier mill database. The good and bad process data was unified into a single dataset. The multivariate method of Partial Least Squares-Discriminant Analysis (PLS-DA) was used to test if the supplier process was operating differently for good and bad runs on the customer's machine. PLS-DA was able to use the process variable data to identify and delineate good and bad supplier product. This paper focuses on communicating a sophisticated statistical analysis method to both supplier and customer mill management, engineering and operations staff.

Identification and Classification of Intermittent Demand Patterns

✱Jing Chen, Temple University, 1810 North 13th Street, 203 Speakman Hall, Philadelphia, PA 19122, *chenjing@temple.edu*; Pallavi Chitturi, Temple University; Mark Gershon, Temple University

Key Words: intermittent demand, lumpy demand, erratic demand, limited demand

Intermittent demand is often defined as random demand with a large proportion of zero values (Silver, 1981). However, the terms 'intermittent demand,' 'lumpy demand,' and 'erratic demand' are used interchangeably in the literature. There is also a widely held misconception that low demand items are, by definition, intermittent in nature. In this paper we show that the term intermittent demand includes at least three distinct sub-groups of demand patterns that we call lumpy, limited, and erratic. We define the three different demand patterns and provide a step-by-step method for classifying items as members of each demand pattern. We provide examples of each demand pattern using inventory data from the U.S. Navy. Forecasting methods to deal with these demand patterns are not finalized or tested. However, we discuss some preliminary strategies for dealing with each distinct demand pattern.

Sequential Analysis for a Binomial Proportion Using an Imprecise Prior Distribution

✱Dennis Oberhelman, University of South Carolina, Moore School of Business, Columbia, SC 29208, *fsu00018@moore.sc.edu*; Kathleen Whitcomb, University of South Carolina

Key Words: sequential sampling, binomial parameter

We describe a procedure for constructing a sequential sampling plan for a binomial parameter, ?, using a class of conjugate priors defined using the 'intervals of measures' method (DeRobertis and Hartigan, 1981). This approach is useful in cases where only partial prior information is available. Upper and lower expected costs for accepting or continuing to sample are computed for each point on the sampling grid, leading to 'upper' and 'lower' sampling plans. A single sampling plan is then constructed to determine whether to 'accept,' 'reject,' or 'continue' for any point on the sampling grid using the 'P-minimality' decision criterion (Troffaes, 2004). An example of a sampling plan using a class of conjugate priors is presented as well as preliminary simulation results comparing expected costs of plans using precise conjugate priors versus classes of conjugate priors.

An Alternative Tolerance Limit for the One-Way Normal Random Effects Model

✱Shun-Yi Chen, Tamkang University, Department of Mathematics, Tamsui, 251 Taiwan, *sychen@mail.tku.edu.tw*

Key Words: Confidence levels, β-content interval, Quantile

An easy computed alternative method of constructing accurate lower tolerance limits for the balanced one-way normal random effects model is derived. Extensive simulation studies indicate that the new procedure is less conservative than that of several current existing methods and gives more accurate coverage rates and smaller standard deviations. Numerical examples are given to illustrate the use of the present procedure.

Deploying Statistical Process Capability Approaches for Assessing Health Care Revenue-Cycle Management

✱Jason Gillikin, Spectrum Health Hosptials, 100 Michigan St NE MC 157, Grand Rapids, MI 49503, *jason.gillikin@spectrum-health.org*

Key Words: hosptial, healthcare, SPC, quality

Much of publicly reported health data depends on information gathered by, and about, a hospital's revenue-cycle operation—registration, patient placement, case management, coding, billing, and denial management. However, most hospitals lack a robust approach to assessing productivity and measuring process/outcome quality for these important functions. This paper summarizes one hospital's adaptation of basic statistical process capability to front-end revenue-cycle operations, with an emphasis on the change of perspective necessary to deploy an SPC model in health care, as well as a summary of major benefits obtained from the change, lessons learned, and recommendations to others in the hospital world. Emphasis is on process improvement and statistical perspectives, and not on statistical procedures.

Supervised Texture Classification Using Vector Autoregressive Models

✱Martin Heller, North Carolina State University, 5125 Powell Townes Way, Raleigh, NC 27606, *maheller@gmail.com*

Key Words: texture classification, random media, vector autoregressive, imaging, fields

We show that for a large class of Markov random fields, the projection of local regions into vector autoregressive (VAR) models makes sense. Of these, there is an equivalence between certain simultaneous autoregressive (SAR) processes and VAR processes. Using the vector autoregressive parameters, one is able to effectively classify disparate textures through supervised learning. In spite of their equivalence, experiments show that one is able to classify textures more accurately using VAR with smaller windows than the SAR models. In addition the elongated rectangular windows required for the VAR estimates allows one to classify texture patterns the SAR model is incapable of classifying with any accuracy.

Effective Choice of Test Boundaries for Sequential Testing Based on Theory of Continued Fractions

✱Yefim H. Michlin, Technion - Israel Institute of Technology, Hashiqma st. 6/19, Nesher, 36812 Israel, *yefim@technion.ac.il*; Genady Y. Grabarnik, IBM T. J. Watson Research Center

Key Words: Sequential testing, Binomial distribution, Continued fractions, Sample number

In an earlier work devoted to sequential testing of binomial processes or processes reducible to ones, we established that the error probabilities of the first and the second kind are incapable of analytical formulation but have a discrete nature. Hence, choosing the optimal testing parameters requires a search for extremes over discrete sets. In our current work we propose a procedure, based on application of the theory of continued fractions, which makes it possible to choose appropriate inter-step intervals in the search for the optimal test boundaries. These intervals have to differ for the accept and reject boundaries, and to depend on the truncation level or on the distribution of the sample number up to the decision.

516 Analysis of Designed Experiments

Section on Physical and Engineering Sciences, Section on Quality and Productivity

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Product and Process Improvement Using Mixture-Process Variable and Robust Optimization Methods

Narinder S. Sahni, Jawaharlal Nehru University; ✱Greg F. Piepel, Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352, *greg.piepel@pnl.gov*; Tormod Naes, MATFORSK

Key Words: Bootstrap, Mixture Experiment, Model Uncertainty, Robust Estimation, Robust Optimization, Split-plot Experiment

A food process was investigated to optimize the raw material (RM) recipe and process variable (PV) levels. Investigations included: (1) optimizing the recipe for discrete combinations of the PVs, (2) optimizing the recipe to be robust to variations in the PVs, (3) optimizing the PV levels to be robust to variations in the recipe, (4) accounting for a cost constraint, and (5) quantifying the uncertainty in the robust, optimal settings resulting from model estimation and model selection uncertainties. A mixture-process variable (MPV) design was constructed and run as a split-plot experiment. Separate models in the RM proportions and PV levels were fit and used to develop combined MPV models. Squared error loss, mean-squared error, and bootstrap methods were applied to develop optimal and robust solutions. Several challenging aspects of these problems and their solutions are discussed.

Relative Effects in Two-Level Experiment Supersaturated or Incomplete Data Sets

✱Linda Trocine, University of Central Florida, c/o Venutek, LLC, P.O. Box 2331, Oviedo, FL 32762-2331, *ltrocine@productivityapex.com*; Linda C. Malone, University of Central Florida

Key Words: incomplete data, design of experiments, data analysis, factor screening, optimization, effect size

We describe a new factor screening analysis method of two-level designed-experiment data sets that are incomplete or are supersaturated. Typically, two-level full factorial designed-experiment data sets are analyzed using effect size estimates which require complete, balanced data sets. F-statistics in an analysis of variance are used for the significant effects. Both analysis methods require a greater number of observations than number of effects being measured (i.e. not saturated). Our method provides a relative measure of the magnitude of the effect that the factors and their interactions have on the response. The effects can then be selected for use in further experimentation and optimization. In one comparative example, the five significant effects from fifteen main and interaction effects screened using effect size estimates, were found in as few as 11 observations with our method.

Discrimination Between Two Linear Regression Models: Design and Inference

✱Rupam R. Pal, University of California, Riverside, 200 W Big Springs Rd, Apt 9, Riverside, CA 92507, *rpal001@student.ucr.edu*; Subir Ghosh, University of California, Riverside

Key Words: Regression Models, Discrimination between Models, Designs, Criterion functions, Fitting, Prediction

We consider the problem of discriminating between the simple linear and quadratic regression models in the presence of the intercept parameter with one explanatory and one response variable. The issue of design selection is discussed with replicated observations at three distinct points under two different allocations for a fixed total number of observations. We also present the discrimination problem between the above mentioned models with no intercept parameter. Several criterion functions are used for the purpose of discrimination. The designs are obtained satisfying the goals of criterion functions for our model discrimination.

Determining Interacting Variables in Random Forest Models Using Fractional Factorial Designs

✱Ewa M. Sztendur, Monash University, Clayton campus, Melbourne, International 3041 Australia, *ewa.sztendur@buseco.monash.edu.au*; Neil T. Diamond, Monash University

Key Words: Random Forests, Variable Importance, Interactions, Fractional Factorial Designs, Plackett-Burman Designs

Random Forests are a powerful classification technique, consisting of a collection of decision trees. One useful feature of Random Forests is the ability to determine the importance of each variable in predicting the outcome. This is done by permuting each variable and computing the prediction accuracy before and after the permutation. This variable importance calculation is similar to a one-factor-at a time experiment, and therefore it provides no information on interactions between variables. In this study, we use regular fractional factorial designs and Plackett-Burman designs to determine which variables to permute. Based on the results of the trials in the experiment, we calculate improved estimates of the individual importance of the variables, and determine which variables interact. The methods are illustrated with a study of student attrition at Monash University.

Industrial Split-Plot Experiments with Non-Normal Data: Go Bayesian

✱Timothy J. Robinson, University of Wyoming, Dept. 3332, Department of Statistics, Laramie, WY 82070, *tjrobin@uwyo.edu*; Christine Anderson-Cook, Los Alamos National Laboratory; Michael Hamada, Los Alamos National Laboratory

Key Words: Response Surface, Bayes, Mixed Model, Generalized Linear Model, Industrial Experiments, Restricted Randomization

In this talk, the analysis of industrial split-plot experiments with non-normal responses is discussed. Specifically, we consider the common situation in which the experimenter is interested in making inferences about characteristics of the response distribution such as quantiles or exceedance probabilities (ex. percentage within specification limits). This talk demonstrates the appeal of Bayesian methods in such situations. An example from film manufacturing is used for illustration and the talk is tailored for those with little or no Bayesian experience who would like to learn more of the utility of these methods.

Object-Oriented Control of Experiments and Information Capacity of Experimental Data

✱Anatoly A. Naumov, Novosibirsk State Technical University, P. Box 158, Novosibirsk, 630091 Russia, *a_a_naumov@mail.ru*

Key Words: Control of experiments, Process approach, Cost expenses, Risks, Design of experiments, Information Capacity

The tasks of experiments control in view of singularities of research object are researched.The problems of exactitude of the soluble tsks, of cost of experiments are considered. The process approach as a basis for construction of control of experimental researches mathematical model tasks is put. It has allowed to consider the tasks from common positions and to solve the tasks of risks estimation. The procedures of an informative capacity of experimental data estimation are researched. So, is shown that it is impossible to estimate an absolute value of informative capacity of the data. The outcomes of researches can be used in practice and will help to explain some problems of practical usage of classical optimal designs of experiments. It is shown that process approach from common positions is a convenient formal basis for optimization of schemes of experimental researches control.

Bootstrap Inference on the Characterization of a Response Surface

✱Robert Parody, Rochester Institute of Technology, 98 Lomb Memorial Drive, Rochester, NY 14623, *rjpeqa@rit.edu*

Key Words: Bootstrap, Simulation, Response Surface, Eigenvalue

Response surface methodology often focuses on finding the levels of some predictor variables that optimize the expected value of a response variable. The eigenvalues of the matrix of pure and mixed quadratic regression coefficients are critical in assessing the nature of higher dimensional second-order response surface models. The estimated values for the eigenvalues will change based on the sample data. Confidence intervals for the eigenvalues could be used to assess this variability. The usual approach is to assume that the error vector is multivariate normal with mean vector 0 and covariance matrix. The approach given in this article will generalize this technique using a nonparametric bootstrap based on a pivot to create approximate confidence intervals for the eigenvalues. This technique will be illustrated with an example and validated with a simulation study.

517 Nonparametric and Semiparametric Inference for Lifetime and Survival Data ●▲

Section on Nonparametric Statistics, Section on Quality and Productivity

Thursday, August 7, 10:30 a.m.–12:20 p.m.

A Nonparametric Model of Quality Adjusted Lifetime Data Analysis with Zero Duration Health States

✱Kaushal K. Mishra, North Carolina State University, 3257 Octavia Street, Raleigh, NC 27606, *kkmishra@gmail.com*; Sujit Ghosh, North Carolina State University

Key Words: Nonparametric Models, Quality Adjusted Lifetime

Clinical trial studies are often conducted in which quality of life is accessed and recorded along with clinically measurable endpoints. Consideration of the quality of life along with the survival time in the statistical analysis can result in better assessment of the drug being tested. Quality adjusted lifetime (QAL) data analysis can serve as an important tool to the medical and patient community. A nonparametric model based on a progressive health state assumption with a data augmentation scheme which provides a nonzero probability to the zero time spent in any health state has been developed. Simulation studies were performed to validate the statistical inference. A real data set was used to illustrate the application of the proposed method and was compared with the parametric models proposed by the authors before.

Semiparametric Efficient Estimation in the Proportional Odds Cure Model

✱Meng Mao, University of California, Davis, Department of Statistics, University of California, Davis, Davis, CA 95616, *mmao@wald.ucdavis.edu*

Key Words: EM algorithm, MM algorithm, Newton-Raphson algorithm, logistic regression, nonparametric maximum

We present a mixture cure model with the survival time of the "uncured" group following the proportional odds property. The EM algorithm is used to locate the nonparametric maximum likelihood estimator (NPMLE). Unlike the proportional hazard mixture model, there is difficulty in the M step triggered by the nonparametric component, so we propose a hybrid algorithm by applying the MM algorithm in the beginning and then switch to the Newton-Raphson algorithm. This hybrid algorithm provides both computational stability and efficiency as demonstrated by simulation studies. Consistency and asymptotic efficiency of the nonparametric maximum likelihood estimator are derived. A case study of leukemia data is conducted to illustrate the proposed procedure, which is also compared to an existing procedure in a simulation study.

New Approaches to Life Testing with a Guaranteed Survival Time

✱Mohammad B. Sepehrifar, The University of Mississippi, Mathematics, The University of Mississippi, Hume Hall 35, P. O. Box 1848, University, MS 38677, *moe@olemiss.edu*

Key Words: U-statistic, Reversed hazard rate, Inactivity time, Classes of life distributions, Exponentiality, Empirical power

Knowing the class to which a life distribution belongs gives us an idea about the aging of the device or system the life distribution represents. The purpose of this talk is to introduce new nonparametric aging classes of life distributions defined by the concept of inactivity time of a unit with a guaranteed survival time. These classes provide a meaningful framework for modeling and studying lifetime data for most physical systems. Also we address the question of testing exponentiality while the alternative is one of these new classes. Several new approaches are discussed. One is based on the empirical distribution function while the others are based on the negative moment inequalities and Laplace transform inequalities.

A Nonparametric Method for Assessment of Interactions in a Multifactor Design with Censored Data

✱M. Hossein Rahbar, The University of Texas Health Sciences Center, P.O. Box 20186 RAS E627, Houston, TX 77030, *Mohammad.H.Rahbar@uth.tmc.edu*; Hooshang Talebi, University of Isfahan; Sangchoon Jeon, Michigan State University

Key Words: Survival Time, Censoring, Multi-factor design, Interactions, Simulations, Regression

We propose a nonparametric method for assessment of interactions in a multi-factor model based on censored data. Specifically, we assess the interaction between a treatment and a covariate on the mean survival time of subjects assigned to combinations of the levels of factors when the effect of treatment is dependent on the covariate. The proposed method allows for varying levels of censorship in the outcome at different levels of the factors. We present the asymptotic distribution of the estimators of the parameters in our model. Through a simulation study we assess the effect of the censoring rate on the standard error of these types of estimators and power of our method with that of using the Cox Proportional Hazard Model. Finally, we demonstrate real life application of our proposed method, by using data from a nursing intervention trial of cancer patients undergoing chemotherapy.

Another Look at the Mean Lifetime Based on Censored and/or Truncated Data

✱Jong S. Kim, Portland State University, Department of Mathematics and Statistics, Portland, OR 97207, *jong@pdx.edu*

Key Words: censored and/or truncated data, mean, nonparametric maximum likelihood estimator, self-consistent estimator

Estimation of the mean lifetime based on censored and/or truncated data is considered. This is an extension of Datta's (2005) work on right-censored data. The usual sample mean cannot be computed due to censoring and/or truncation. A commonly used approach is to compute the area under the nonparametric maximum likelihood estimator of the survivor function. Here, we show that a seemingly different approach -computing the usual sample mean after imputation- is equal to the former as long as the survivor function estimator is self-consistent. Several examples are provided.

Isotonic Estimation of Monotone Hazard Ratios for a Proportional Hazards Model

✱Marek Ancukiewicz, Massachusetts General Hospital/Harvard Medical School, Department of Radiation Oncology, 55 Fruit Street, Boston, MA 02114, *msa@biostat.mgh.harvard.edu*

Key Words: order-restricted inference, isotonic regression, proportional hazards, survival analysis, nonparametric hazard estimation

I consider proportional hazards model, with the hazard of an event proportional to an arbitrary monotone function f(z) (hazard ratio function) of a continuous or ordinal covariate z. I propose a method of estimation of the function f() at observed distinct covariate values z1,…,zk such that the estimate preserves order of z1,…,zk and maximizes the partial likelihood. No other assumptions are made (except for some normalization of f()). The estimate is obtained by extending isotonic regression to survival data and exploiting the log-convexity of the partial likelihood function of f(z1),..,f(zk). The method can be used (e.g., for graphical display of survival data and checking the correctness of a parametric specification of the hazard ratio function). I consider example medical applications and present some extensions of the method (e.g., to multiple and time-dependent covariates).

A Mass Redistribution Algorithm for Right-Censored and Left-Truncated Time-to-Event Data

✱Xu Zhang, Georgia State University, 30 Pryor St, COE 750, Atlanta, GA 30303, *matxxz@langate.gsu.edu*; Mei-Jie Zhang, Medical College of Wisconsin; Jason P. Fine, University of Wisconsin-Madison

Key Words: Cumulative incidence function, Subdistribution hazard, Truncation

Failure times are often right-censored and left-truncated. In this paper we give a mass redistribution algorithm for right-censored and/or left-truncated failure time data. We show that this algorithm yields the Kaplan-Meier estimator of the survival probability. One application of this algorithm in modeling the subdistribution hazard for competing risks data is studied. We give a product-limit estimator of the cumulative incidence function via modeling the subdistribution hazard. We show by induction that this product-limit estimator is identical to the left-truncated version of Aalen-Johansen (1978) estimator for the cumulative incidence function.

518 Improving the Quality of Survey Data

Section on Government Statistics, Section on Survey Research Methods

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Assessing the Effect of Calibration on Nonresponse Bias in the 2005 Agricultural Resource Management Survey Phase III Sample Using Census 2002 Data

✱Morgan S. Earp, National Agricultural Statistics Service, 3251 Old Lee Highway, Fairfax, VA 22030, *morgan_earp@nass.usda.gov*

Key Words: Nonresponse Bias, Response Rate, Calibration Weights, Mean Relative Bias

The Agricultural Resource Management Survey (ARMS) is conducted by the USDA and collects detailed economic data from US producers. As with most surveys, ARMS suffers from unit nonresponse (70.5% in the 2005) with the potential to introduce nonresponse bias. Nonresponse bias was assessed by matching records sampled for the ARMS with those from the 2002 Census. Mean relative bias was assessed for 17 variables by comparing estimates based on census data for all ARMS cases (respondents and nonrespondents) versus ARMS respondents, using uncalibrated sampling weights and calibrated weights. Nine of the 17 had significant bias using the ARMS base weights. The ARMS calibration weights reduced the bias so that it was no longer significantly different from zero in 90% of the study variables. This suggests that calibration is an effective tool in reducing nonresponse bias to acceptable levels.

Adjustments for Mode Effect Bias for the Canadian Community Health Survey

✱Andrew Quigley, Statistics Canada, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6 Canada, *andrew.quigley@statcan.ca*

Key Words: mode effect, Canadian Community Health Survey

The Canadian Community Health Survey is a cross-sectional survey designed to produce nationally comparable estimates of health risk factors, health status and health care services. Annually, 65,000 respondents are collected through a mix of computer-assisted telephone interviews (CATI) and computer-assisted personal interviews (CAPI). To increase the precision of estimates, provinces can provide extra funds to increase the sample size for their area. Almost all of these 'buy-in' units use the more cost-effective CATI, thus leading to a larger than average ratio of CATI to CAPI interviews. For variables affected by the mode of collection, this larger ratio leads to problems in comparing with past results or results from other areas. This paper identifies the presence of a mode effect for some variables and suggests a method to overcome the problem using a reweighting technique.

Methods To Promote Deeper Processing on Survey Questions

✱Emily McFarlane, RTI International, 3040 Cornwallis Rd, Durham, NC 27709, *emcfarlane@rti.org*; Robert Steele, RTI International; Laura Burns, RTI International; Daniel Pratt, RTI International

Key Words: Survey methodology, forced-choice

Research has shown that forced-choice questions promote deeper processing and less satisficing, resulting in greater numbers of responses being selected (Smyth et al., 2005). The authors tested whether deeper processing can also occur when this format is applied to other types of questions. In a survey for the National Science Foundation, respondents are asked to update a list of science and engineering departments offered at their school. The authors tested two versions of presenting the listing task to respondents. In one group, respondents were asked to update their listing as needed and then check a box at the bottom of the page to confirm they had updated their listing. In the other group, respondents confirmed or deleted each unit individually - a forced-choice format. The authors hypothesized that the forced-choice format would promote deeper processing in respondents.

Improving Educational Data Quality for Growth Model Analysis

✱Kening Wang, University of Arkansas, 302 WAAX, 340 N.West Ave,, Fayetteville, AR 72701, *kxw06@uark.edu*; Sean Mulvenon, University of Arkansas; Charles Stegman, University of Arkansas - NORMES; Yanling Xia, University of Arkansas

Key Words: Data Quality, Growth Model, Longitudinal Database

U.S. Department of Education initiated pilot growth model program in 2007 to implement Adequate Yearly Progress (AYP) evaluations for schools under No Child Left Behind (NCLB) legislation. Measuring whether a student has improved academically is an important component of growth models. Repeated measures or randomized block designs are traditional statistical methods to determine if a student has improved. Some new methods, such as Performance Growth Index (PGI), have been proposed recently. Accuracy of growth model is highly depended on the quality of longitudinal educational database, because the growth of individual student is calculated in successive grades or years. A lot of factors affect tracking of student's records. This paper will discuss our experience with longitudinal data quality problems and providing an overview of the main solution approaches.

Use of a Listed Sample To Supplement and Improve the Accuracy of a Probability Sample

✱Madhumita (Bonnie) Ghosh-Dastidar, RAND Corporation, 37027, *bonnieg@rand.org*; Marc N. Elliott, RAND Corporation; Amelia Haviland, RAND Corporation; Lynn Karoly, RAND Corporation

Key Words: composite estimator, design effect, listed sample, minimum MSE estimator, random digit dial

Random digit dialing can be a very costly methodology for a rare population, but less costly listed samples are not representative and can induce bias. When the list source has less than 30% coverage of the population, the maximum contribution of the listed sample to precision may be severely limited. In a telephone survey of child care for 3- 4-year-old children in California, we conducted parallel RDD and listed samples and parallel weighted estimation within each sample. We combined these parallel estimates to produce minimum MSE composite estimates for parameters of interest, with weights inversely proportionate to the estimated MSE of each estimate. In particular, we estimated the bias of the listed estimate using the RDD estimate as a gold standard. This approach allows the listed sample to contribute differently to different parameter estimates as a function of estimated bias.

Fuzzy Matching of Federal Science and Technology Program Data

✱Gary W. Anderson, National Institute of Standards and Technology, 100 Bureau Dr, MS 106, Gaithersburg, MD 20899, *gary.anderson@nist.gov*

Key Words: Innovation, R&D, Evaluation

Innovation has long been credited as a leading source of economic strength and vitality in the United States. Existing data sources can be leveraged to develop better measurement of innovative activities and a better understanding of the innovation process. This paper reports on the results of a data linking project that matches nearly 65,000 participants in four federal science and technology (S&T) programs with Census microdata. This paper uses fuzzy-matching techniques to link the data collected through these programs with establishment and company level data in the Census Bureau's Longitudinal Business Data (LBD) and Business Register (BR) and Survey of Industrial Research and Development (SIRD). This matching will the census frame and lead to a better understanding of the innovation process and federal program impact.

Maximum Likelihood Estimation for Categorical Data with an Almost-Missing-at-Random Mechanism

✱Yan Zhou, The University of Michigan, Dept. of Biostatistics, Ann Arbor, MI 48105, *yzhouz@umich.edu*; Roderick J. Little, The University of Michigan; John Kalbfleisch, The University of Michigan

Key Words: missing, nonignorable, maximum likelihood estimate

Missing values are common issues in empirical studies. The performance of the methods analyzing missing data strongly depends on the missing-data mechanism. If missing mechanism is ignorable, the likelihood based inferences will only depend on the observed data. For two variables X and Y where both of X and Y are missing, however, missing-data mechanism may not be ignorable, since the missingness of Y can possibly depend on the values of X which are missing. We propose a nonignorable missing-data mechanism, in which X is MCAR and Y is MAR given the value and missingness of X. The non-iterative maximum likelihood estimates exist and data are excluded for estimating certain parameters. Extensions of this type of mechanism will be also discussed.

519 The Internet, Telecommunications, and Migration

Business and Economics Statistics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

E-Commerce Online Experiments: Design and Analysis Challenges

✱Victor N. Morin, Ecolab Research Center, 655 Lone Oak Drive, ESC-F4316, Eagan, MN 55121-1560, *victor.morin@ecolab.com*; Jonathan Francis, Microsoft

Key Words: e-commerce, experiments, web, online

Online experiments are conducted on e-commerce web sites to optimize site design to increase conversion rates. There are many interesting and challenging statistical issues involved in the design and analysis of e-commerce online experiments, including sample sizes, sample assignment to treatment groups, and the statistical analysis of the large data sets obtained, data that is often nonnormal and containing outliers. This paper will review the key challenges, present examples of data obtained from online experiments, and review how two major e-commerce sites have approached these challenges. The challenges involved in e-commerce online experiments have had little coverage in the statistical literature. It is hoped a review of the challenges and examples of actual experiment results will increase interest in developing robust approaches to address these challenges.

Functional Data Analysis for Bid Arrival Times in eBay Auctions

✱Shuang Wu, University of California, Davis, 5000 Orchard Park Cir 5725, DAVIS, CA 95616, *swu@wald.ucdavis.edu*; Hans G. Müller, University of California, Davis

Key Words: Density function, eBay auction, Functional data analysis, Intensity function

Consider the bid arrival times for online eBay auctions where one observes a relatively small number of irregularly spaced bids for each auction. We propose to model bid arrivals by postulating a random rate function as realization of an underlying stochastic process for each auction. Focusing on the random density function that reflects the shape of the rate function, a functional data analysis approach is developed by estimating mean and covariance structure of these random density functions, followed by the estimation of eigenfunctions and functional principal component scores to obtain fitted density and also rate functions for individual auctions. Theoretical results include asymptotic consistency for the components of the model. The method is illustrated with a study of bid arrival times for 156 Palm M515 PDA auctions from eBay.

An Investigation of Virtual Worlds Adoption: A Research Framework and Empirical Study

✱Guangying Hua, Bentley College, 175 Forest Street, Waltham, MA 02452, *ghua@bentley.edu*; Dominique Haughton, Bentley College

Key Words: Structural Equation Modeling, Survey research, Virtual Worlds, Technology Acceptance Model, Diffusion of innovation

A virtual world is an Internet-based simulated environment which intends for its users to inhabit and interact via avatars. As a new information technology innovation, it has developed so fast with the economic success for organizations and individuals. Very limited research has been done to investigate the virtual world so far. This paper presents a research framework to investigate the acceptance of the virtual world from individual users' perspectives. In order to cater to the special features of the virtual world, a framework based on the technology acceptance model and diffusion of innovation is presented from a synthesis of the extant literature. An empirical study is employed to test each hypothesis. The results show useful implications.

Tracking Changes in Rates of Return to Revenues and Expenses with Prices Set According to Target Rates in Regulated Telecommunications Companies

✱Martin B. Selzer, Chatham Decision Sciences, 72 Hedges Avenue, Chatham, NJ 07928, *mselzer@att.net*

Key Words: Financial Measurement, Rate of return tracking, Regulated industries, Telecommunications, Reported data changes

The National Exchange Carrier Association (NECA) is a not-for-profit association of local rural telephone companies. NECA sets prices for network access that the local companies charge long distance companies. These prices are set targeting a rate of return on investment authorized by the FCC. As companies report revenues and expenses, actual rates of return are realized. The FCC rules allow companies to revise these revenues and expenses monthly for 24 months after the initial reporting of values. This results in a monthly fluctuation of the rate of return earned by companies. To understand the source of these fluctuations, NECA traces the rate of return basis points changes to differences in reported revenues and expenses. This paper describes the measurement methodology used to relate changes in revenues and expenses to rate of return fluctuations for companies in NECA.

Tolls, Exchange Rates, and Northbound International Bridge Traffic from Ciudad Juarez to El Paso

✱Angel L. Molina, Jr., University of Texas at El Paso, 500 West University Avenue, UTEP Department of Economics and Finance, El Paso, TX 79968, *amolina@utep.edu*

Key Words: Bridge Traffic, Tolls, Applied Econometrics, Mexico Border

Increased economic activity plus demographic expansion have combined to raise cross-border traffic over the international bridges that connect El Paso, Texas, and Ciudad Juarez, Chihuahua, an area commonly referred to as the borderplex. Previous econometric analyses examine southbound international bridge crossings within this region. This effort attempts to partially fill a gap in the existing literature by analyzing the impacts of toll variations and international business cycle fluctuations on northbound international traffic over a 16-year period. Parameter estimation is conducted using a linear transfer function ARIMA methodology. Price elasticities of demand are similar to those reported for other regional economies, but out-of-sample forecasting results are mixed.

The 'Melting Pot': A Success Story?

✱Edna Schechtman, Ben-Gurion University of the Negev, Reger st., Beer Sheva, 84105 Isreal, *ednas@bgu.ac.il*; Shlomo Yitzhaki, Central Bureau of Statistics/Hebrew University

Key Words: Gini, Classification, Stratification, Migration

The "melting pot" policy intends to create a new generation of migrants that will assimilate with the native population. We examine the success of this policy among second generation Israelis, in terms of income. Is the second generation more similar to the natives, or to their own parents? The methodology used is ANoGI - analysis of Gini - which is similar to ANOVA except that it offers an additional parameter - stratification between groups. The methodology answers the general question: which classifying variable divides the population into subgroups in a better way? The second-generation Israelis were classified as 1. Israelis, and 2. Belonging to the country of origin (of their parents). We find that the latter classification is more stratified than the former. This means that the melting pot policy did not achieve its declared targets.

Traditional Teaching Giving Way to New Techniques

✱Mammo Woldie, Texas Southern University, Dept of Business Administration, 3100 Cleburne, Houston, TX 77004, *woldie_mx@tsu.edu*

Key Words: chalk and board, online, semi-online

In addition to the usual chalk and board approach, nowadays several people use online or semi-online approach in teaching basic statistics. An attempt will be made to examine how much successful these techniques have become in helping students learn.

520 Statistical Methods for QTL Mapping and Tiling Arrays

Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Bayesian Mixture Structural Equation Modeling in Multiple-Trait QTL Mapping

✱Xiaojuan Mi, University of Nebraska, Department of Statistics, 340 Hardin Hall North,, East Campus, Lincoln, 68583-0963, *xjmixu@yahoo.com*; Kent M. Eskridge , University of Nebraska; Dong Wang, University of Nebraska-Lincoln

Key Words: Bayesian, QTL mapping, mixture SEM, multiple-trait

Research on mapping quantitative trait loci (QTL) often results in data on a number of traits that have well established causal relationships. In this paper, a mixture SEM with a Bayesian approach for QTL mapping is developed to take into account the causal relationships among traits and QTL. Since the posterior distribution of the mixture SEM is very complicated, a latent allocation variable is introduced which separates samples into multiple subsamples. The analysis can be carried out separately for each sample. Under the Bayesian framework, the conjugate priors are chosen for the parameters. The posterior distributions of parameters are estimated using Markov Chain Monte Carlo methods (Gibbs sampler algorithm). The performance of the proposed method is evaluated by simulation and with data from a recombinant inbred lines (RILs) population from a wheat genetics experiment.

Computationally Efficient Estimation of False Discovery Rate in eQTL Studies Using Sequential Permutation P-Values

✱Timothy Bancroft, Iowa State University, Ames, IA 50011, *timmyb@iastate.edu*; Dan Nettleton, Iowa State University

Key Words: eQTL mapping, Permutation testing, False discovery rate, Multiple testing, Sequential analysis

This talk describes a computationally efficient method for permutation testing in quantitative trail loci (QTL) mapping studies involving multiple traits. For each trait, the sequential procedure of Besag and Clifford (1991) is applied to the permutation distribution of the maximum test statistic across all tested loci. The resulting permutation p-value can be used to test the null hypothesis of no association between the trait and the identified locus while controlling the genomewise type I error rate for each trait. To account for multiple testing across thousands of traits - as encountered in expression

QTL (eQTL) studies - we develop an extension of the approach of Nettleton et al. (2006) and apply it to the collection of sequential permutation p-values to obtain approximate control of the false discovery rate. We demonstrate the method using data from an eQTL study.

Normalizing and Peak Finding for Tiling Microarrays with Application in Epigenomics

✱W. Evan Johnson, Brigham Young University, Department of Statistics, TMCB 230, Provo, UT 84602, *evan@stat.byu.edu*

Key Words: tiling microarray, epigenomics, normalization

This talk will present a novel G-C based algorithm for normalizing two-color tiling microarray data. This method is very flexible and ideal for removing outliers in data caused by differences in the binding affinity of microarray probes. Additionally, a method high-level analysis method is presented to identify interesting genomic features. The method has been shown to be very sensitive and can therefore detect regions with only slight differences between the arrays. An application from epigenomics is given based on the DNA methylation pattern in model organism.

Analysis of Arabidopsis Genome Tiling Arrays To Detect Nuclear Pre-mRNA Processing Events and Differential Gene Expression

✱Ann Hess, Colorado State University, 1877 - Department of Statistics, Fort Collins, CO 80523, *hess@stat.colostate.edu*; Amber Hackstadt, Colorado State University; A.S.N. Reddy, Colorado State University; Arthur Hunt, University of Kentucky

Key Words: microarrays, alternative splicing

Nuclear pre-mRNA processing events, especially alternative splicing and alternative polyadenylation of pre-mRNAs, are important regulators of gene expression. These processes generate multiple transcripts from a single gene and thereby contribute greatly to transcriptome complexity and gene regulation in eukaryotic organisms. Genome tiling arrays offer an unprecedented opportunity to investigate these processes on a global scale. Here we discuss some progress and challenges in analyzing Arabidopsis tiling array data to detect alternative splicing as well as differential gene expression. Analysis of tiling array data for these purposes is largely unexplored and requires novel approaches. Our approach uses the mixed model framework.

Variable Selection Method for Quantitative Trait Analysis Based on Parallel Genetic Algorithm

✱Siuli Mukhopadhyay, Indian Institute of Technology Bombay, Department of Mathematics, Powai, Mumbai, International 400076 India, *siuli@math.iitb.ac.in*; Varghese George, Medical College of Georgia; Hongyan Xu, Medical College of Georgia

Key Words: Haplotype, Tagging SNPs, Combinatorial Optimization, Bayesian Variable Selection

Most of the traits in complex human diseases are quantitative in nature. Generally, these quantitative traits are jointly affected by multiple genetic and environmental factors. The genetic basis of most of these complex traits is unknown and there is strong interest in genetic analysis of the quantitative traits. It is the purpose of these analyses to identify the relevant genetic and environmental factors, which are important for understanding the etiology and treatment of human diseases. In this study, we use parallel genetic algorithm (PGA) and variable selection to develop strategies for identifying the important genetic and environmental factors. Our method can take account of both multiple markers across the genome and the environmental factors. Using simulated and real examples we show that PGA is able to choose the variables correctly and also easy to use computationally.

521 Methods for Integration of Genomic Data

Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

A Meta-Analysis Approach for Gene Association Network Reconstruction

✱YounJeong Choi, University of Wisconsin-Madison, Department of Statistics, 1300 University Ave, Madison, WI 53706, *ychoi@stat.wisc.edu*; Christina M. Kendziorski, University of Wisconsin-Madison; Alan D. Attie, University of Wisconsin-Madison; Mark P. Keller, University of Wisconsin-Madison

Key Words: Gene association network, Meta-analysis, Gene expression, Microarray

Gene association networks (GANs) are important in studies of complex traits as they provide a simple, yet effective, summary of relationships among genes. Most GAN reconstruction algorithms link genes having high marginal or partial correlation. However, we find that the presence of either correlation is neither necessary nor sufficient among genes in known pathways. To address the problem, we have developed an approach for GAN identification that does not require strong correlations. It utilizes data from multiple studies in multiple species, provides FDR control, and is more powerful than competing approaches. An analysis of gene groups defined from Gene Ontology demonstrates the ability of the approach to identify GANs important in diabetes.

Accuracy of 'Network Filtering' for Detecting Perturbations in Large, Sparse Networks

✱Shu Yang, Boston University, 111 Cummington street, Boston, MA 02215, *shuyang@math.bu.edu*; Eric Kolaczyk, Boston University

Key Words: Sparse high dimensional network, Lasso, Network, Filtering

Network filtering has been proposed recently to detect the effects of certain external perturbations on the interacting network members. However, with large networks, the goal of detection seems a priori difficult to achieve, especially when the number of observations available is much smaller than that of variables describing the effects of the underlying network. Assuming the network structure possesses a certain sparsity property, we provide a formal characterization of our ability to detect the external effects, using a network filtering system that combines Lasso regression in a sparse simultaneous equation model with simple residual analysis. We explore the implications of the conditions underlying our characterization, in the context of various network topologies, and illustrate our method using both simulated data and a problem of drug target detection from DNA microarray data.

A Statistical Framework for Integrating Different Microarray Data Sets in Differential Expression Analysis

✱Yinglei Lai, The George Washington University, 2140 Pennsylvania Avenue, N.W., Washington, D.C., DC 20052, *ylai@gwu.edu*

Key Words: microarray, integration, differential expression, concordance, mixture model

We expect to achieve more efficient detection of differential expression if an efficient statistical method can be developed to integrate different microarray data sets for the same study. We propose to first evaluate the genome-wide p-values for each individual data set and then transform the p-values

into their corresponding z-scores to achieve efficient normal mixture modeling. Before the data integration, it is necessary to perform the concordance/discordance tests. Based on the test results, different subsequent actions are suggested. The data integration can be considered if we can reject the hypothesis of complete discordance; otherwise, the data integration is discouraged. We develop a bivariate normal mixture model based method to integrate different lists of genome-wide z-scores. Genes can be prioritized by their probabilities of being concordantly differentially expressed.

Genome-Wide Associations Between SNP Genotypes and Gene Expression, Detected by Both Frequentist and Bayesian Approaches

✱Amrita Ray, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 977R225A, Berkeley, CA 94720, *aray@lbl.gov*; Paul T. Spellman, Lawrence Berkeley National Laboratory; Terence P. Speed, University of California, Berkeley

Key Words: single nucleotide polymorphism, gene expression, bayes factors, association analysis, multiple testing correction

This study performs association analysis of human gene expression with dense sets of single-nucleotide polymorphism (SNP) markers. The Affymetrix Exon 1.0 ST expression profile and Affymetrix SNP6.0 genotype data of 200 GBM tumors are obtained from the Cancer Genome Atlas database. The 'good quality' SNPs are selected, and those within a 50 kilobases region flanking either side of a gene are considered. For each gene, association of expression level to SNPs is tested using a linear regression model, and a modified t statistic in bayesian framework is developed. The raw p-values are corrected by Benjamini-Hochberg multiple testing procedure, and region bayes factors are computed. Preliminary results show 2% of the genes are associated with the SNPs (p-value of < 10-3), with 30% of these pvalue < 10-6. Thus association studies with dense SNPs can identify determinants of gene expression.

A Joint Model for Nucleosome DNA Sequence and Micrococcal Nuclease

✱Liqun Xi, Northwestern University, 2006 Sheridan Road, Dept of Statistics, Evanston, IL 60208-4070, *lxi700@northwestern.edu*; Yvonne Fondufe-Mittendorf, Northwestern University; Jonathan Widom, Northwestern University; Ji-Ping Wang, Northwestern University

Key Words: alignment, joint statistical model, linker DNA, micrococcal nuclease, nucleosome DNA, yeast

High throughput mapping of nucleosome DNA sequences has shown that special motifs are conserved but weakly at certain positions of nucleosomal region. However such nucleosome core sequences come as truncated or with some extra linker DNAs attached to the core in either end, largely owing to the micrococcal nuclease specificity. Aligning the nucleosome sequences is exceedingly challenging due to interference of nuclease signals. We propose a joint statistical model that accounts for nucelosome core signals and nuclease specificity simultaneously. We show this approach improves the alignment while discovering the nuclease behavior in digesting chromatin fiber. The approach is applied to more than 50,000 nucleosome sequences obtained from yeast with variable lengths from 142 to 152 bp.

Nonparametric Meta-Analysis for Identifying Signature Genes in the Integration of Multiple Genomic Studies

✱Jia Li, University of Pittsburgh, 735 S Millvale Avenue, Department of Biostatistics, Pittsburgh, PA 15213, *jil53@pitt.edu*; Geroge C. Tseng, University of Pittsburgh

Key Words: Meta analysis, Microarray, Optimal weighted, Maximum p-value

For detection of differentially expressed genes in the meta analyses of different types of microarray data, most of the current efforts are focused on comparing and evaluating gene lists obtained from each individual dataset. The statistical framework is often not rigorously formulated and a real sense of information integration is rarely performed. We tackle two often asked biological questions: "Which genes are significant in one or more data sets?" and "Which genes are significant in all data sets?" We illustrated two statistical hypotheses and proposed an optimal weighted statistic and a maximum p-value statistic for the two questions, respectively. The proposed test statistic is shown to be admissible. And we further show the advantage of our proposed test procedures over existing methods by power comparison, simulation study and real data analyses.

522 Statistical Issues Related to Safety Data ●▲

Biopharmaceutical Section, Biometrics Section

Thursday, August 7, 10:30 a.m.–12:20 p.m.

Safety Findings in Clinical Trials: Are They Real, or Just Coincidental?

✱Mani Lakshminarayanan, Merck & Co., Inc, *Mani_Lakshminarayanan@merck.com*; Amarjot Kaur, Merck & Co., Inc

Key Words: Early Phase trials, Correlated endpoints, Family-wise Error Rate, False discovery rate, Coincidence, Adverse Events

In all phases of clinical trials, the underlying risk of an investigational drug is determined based on multiple type/number of safety endpoints including adverse events, laboratory data, vital signs, electrocardiogram and others. Due to the inherent multiplicity in safety data, inferential procedures (if employed) tend to control false positive rates using procedures based on p-values. Most of these procedures are designed to control either the family-wise error rate (FWER) or the false discovery rate (FDR). The p-plot method is useful in estimating the unknown number of true null hypotheses out of all hypotheses that are involved in the multiplicity. In this paper, we will examine the impact of small samples (Phase-II trials) as well as the correlation between safety endpoints on various measures that are outlined above. Simulated data will be used to compare the results.

A New Approach To Monitor the Safety Risk in Clinical Trials

✱Chenxiong (Charles) Le, MedImmune, Inc., 1 MedImmune Way, Gaithersburg, MD 20878, *lec@medimmune.com*; Qing Liu, Johnson & Johnson Pharmaceutical R&D, LLC

Key Words: Clinical trials, sequential probability ratio test, dose escalation, cardiovascular

Monitoring safety risk is essential in clinical trials to protect patients and to reach decisions on dose selection. The traditional sequential probability ratio test (SPRT) is appropriate for short term safety endpoints, but it can be problematic when safety risk is monitored over a longer period of clinical evaluations. A new method, based on a reverse sampling, is proposed to monitor the safety risk, and the characteristics are investigated. This method is illustrated in an example in a large cardiovascular dose escalation study. The boundary of the SPRT is also investigated when the sample size is small.

Analysis of Vaccine Adverse Event Count Data with Missing Safety Follow-Up: A Multiple Imputation Method

✱Xiaoming Li, Merck Research Laboratories, UG1CD-44, 351 N. Sumneytown Pike, North Wales, PA 19454, *xiaoming_li@merck.com*; Jin Xu, Merck & Co., Inc.; Ivan S.F. Chan, Merck Research Laboratories

Key Words: missing data, safety follow-up, multiple imputation, proportion, incidence rate

Consider a trial in which subjects receive either vaccine or placebo and are followed for serious adverse events (SAE) for 6 months post vaccination. Some subjects may report correlated multiple SAEs and other subjects may not complete the follow-up period, resulting in missing data. One study objective is to compare the relative risk of SAE (vaccine/placebo), and the variables of interest are the proportion of subjects reporting SAE and the incidence of SAE (accounting for follow-up time). A widely used method is the Miettinen and Nurminen method for comparison of two rates. However, this method does not account for missing data and the results may be biased if the data is missing at random. We propose a propensity score-based multiple imputation method as a sensitivity analysis. By simulation, different imputation methods are investigated and compared with the existing method(s).

Using the Z-Test To Sort Association Rules in Large Spontaneous Reporting Databases

✱Chandra S. Thames, Bracy Analytics Inc., 2300 Lakeview Parkway, Suite 700, Alpharetta, GA 30004, *chandra@bracy-analytics.com*

Key Words: association rules, data mining, vaccine safety, drug safety

For many years the focus in 'mining' association rules in large databases has been on the development of more efficient algorithms for finding such rules. However, in recent years research has also focused on statistical techniques for sorting the rules. We recommend sorting the rules by the z-test of the lift instead of the adjusted lift by the Gamma Poisson Shrinkage (GPS) estimator. The GPS estimator is currently the only method used for searching the FDA's Spontaneous Reporting System database and the CDC's Vaccine Adverse Events Spontaneous Reporting System for 'interesting' drug-event combinations. The mean and variance of the lift are derived in order to perform the z-test. The GPS was developed to 'shrink' the lift, adjusting for small counts. However, those rules with low expected counts still dominate the top after applying GPS.

Parametric Parsimonious Markov Mixtures for Safety Data Analysis

✱Xiaodong Wang, sanofi-aventis, 200 Crossing Blvd, P.O.Box 6890, BX2-409A, Bridgewater, NJ 08807, *sheldon.wang@sanofi-aventis.com*; Hanxiang Peng, University of Mississippi

Key Words: complete monotonicity, Poisson regression

In many clinical studies, safety endpoints such as adverse events are binary data, which are often correlated within each patient. The drug safety analysis is very important to assess the medication safety in drug development. An appropriate statistical model is critical in estimating the safety event probability. In this paper, we propose a new class of parametric parsimonious distributions using Markov chain with completely monotonic functions, and a model selection procedure to analyze the overdispersion safety data. A real case will be discussed to illustrate the procedure. Comparison is made with other models such as Poisson regression model.

Permutation Test Approach To Compare Two Groups of Censored Lognormal Data with Multiple Detection Limits

✱Wei Zhong, ICON Clinical Research, 95134, *zhongw@iconus.com*; Jeffrey Welge, University of Cincinnati; Linda Levin, University of Cincinnati; Paul Succop, University of Cincinnati; Rakesh Shukla, University of Cincinnati

Key Words: Lognormal, Permutation Test, MLE-EM, Type I Error Rate, Power

Censored lognormal data with multiple lower detection limits frequently arise in safety laboratory assessment. A common practice to compare the means between two groups of such data is to perform the two sample t-test on the log-transformed outcome and to report the resulting P-values for the null hypothesis based on the original data. The problem with such a practice is that it essentially tests the equality between the two group medians. The permutation test approach with the mean estimation obtained from the MLE-EM algorithm was studied in this paper. The proposed method was evaluated through a series of simulation studies and its performance was compared to the two sample t-test and Z-score test. Simulation results consistently showed that the proposed approach provides the minimum Type I error rates and maximum power.

Continuing Education Courses

JSM
2008

August 2, 2008

CE_01C 8:30 a.m.–5:00 p.m.

Generalized Linear Mixed Models: Theory and Applications

ASA

Instructor(s): Oliver Schabenberger, SAS Institute Inc.; Walter Stroup, University of Nebraska-Lincoln

This two-day course is for those who want to learn about the theory and application of generalized linear mixed models across disciplines from a non-Bayesian perspective. Each day comprises theory and application components with numerous examples. The material is presented at an applied level, accessible to participants with training in linear statistical models and previous exposure to linear mixed models. On the first day, we will cover classes of mixed models and how their features are made manifest in today's likelihood-based estimation methods. We will make the connection between linear models, generalized linear models, linear mixed models, and generalized linear mixed models (GLMM) in terms of model formulation, distributional properties, and approaches to estimation. Participants will learn that GLMMs are an encompassing family and understand the differences and similarities in approaches to estimation and inference within the family. We will discuss overarching issues that confront analysts who work with correlated, non-normal data, such as overdispersion, the marginal and conditional models, and model diagnostics. During the second day, we will focus on application areas for GLMMs and examples; supporting theory will be introduced as needed. Focus areas will include modeling of rates and proportions, modeling of regular and zero-inflated counts, mixed model smoothing, the computation of power and sample size, and inferential tasks with and without adjustments. Computations will be based on the mixed model tools in SAS/STAT software.

CE_02C 8:30 a.m.–5:00 p.m.

Genetic and Microarray Data Analysis

ASA

Instructor(s): Russ Wolfinger, SAS Institute Inc.; Carl D. Langefeld, Wake Forest University Health Sciences

This course is for statisticians who wish to learn about statistical genetics, microarray data analysis, and prediction with genomic biomarkers. Course content will be at the intermediate level. It time permits, we will cover topics such as copy number, exon arrays, ChIP-on-Chip, and eQTL. There will be a mixture of theory and practical examples. JMP Genomics software and custom scripts will be used for illustration.

CE_03C 8:30 a.m.–5:00 p.m.

Optimal Experimental Designs

Section on Physical and Engineering Sciences

Instructor(s): Alexander N. Donev, University of Manchester; Randy Tobias, SAS Institute Inc.

Optimal design for the practitioner is often discussed as a "black box," shying away from the theory. On the contrary, the premise for this course is that powerful practical approaches for assessing the properties of standard designs and of finding good designs in nonstandard situations result from familiarity with the theory of optimal experimental design. We will start by covering fundamental theory, including forms of the General Equivalence Theorem that are central to algorithms for the construction of optimal designs. These ideas will be illustrated with standard designs for response surface models. We will move on to common nonstandard problems in design for response surfaces, such as blocking, finding designs over irregular regions, and mixture designs. We will also discuss the augmentation of designs and designs for checking the adequacy of models. Many models in chemistry and the pharmaceutical industry are nonlinear in the parameters. Optimal designs for these models depend on prior information about the parameters, which may be available in the form of a prior distribution. We will show how this information may be used to provide good designs.

CE_04C 8:30 a.m.–5:00 p.m.

Regression Modeling Strategies

ASA

Instructor(s): Frank E. Harrell, Jr., Vanderbilt University

All standard regression models have assumptions that must be verified for the model to have power to test hypotheses and predict accurately. Of the principal assumptions, this course will emphasize methods for assessing and satisfying linearity and additivity. Practical but powerful tools will be presented for validating model assumptions and presenting model results. This course provides methods for estimating the shape of the relationship between predictors and response by augmenting the design matrix using restricted cubic splines. Even when assumptions are satisfied, over fitting can ruin a model's predictive ability for future observations. Methods for data reduction will be introduced, methods of model validation will be covered, and auxiliary topics such as modeling interaction surfaces, efficiently utilizing partial covariable data by using multiple imputation, variable selection, overly influential observations, collinearity, and shrinkage will be discussed. The methods covered will apply to almost any regression model, including ordinary least squares, logistic regression models, and survival models.

CE_05C 8:30 a.m.–5:00 p.m.

Hot Topics in Clinical Trials

Section on Teaching Statistics in the Health Sciences, Boston Chapter

Instructor(s): Scott R. Evans, Harvard University; Lee-Jen Wei, Harvard University; Lu Tian, Northwestern University; Lingling Li, Harvard Medical School

We will address several hot-topic areas in clinical trials, including the use of prediction to identify biomarkers, meta-analysis of rare safety events, data monitoring committees, data monitoring using prediction, noninferiority studies, causal inference, benefit:risk assessment, and bridging studies. We will present motivating examples and discuss standard and novel approaches to analyses.

CE_06C 8:30 a.m.–5:00 p.m.

Successful Data Mining in Practice

ASA

Instructor(s): Richard D. De Veaux, Williams College

This course will introduce data mining, which is the exploration and analysis of large data sets by automatic or semiautomatic means with the purpose of discovering meaningful patterns. The knowledge learned from these patterns can be used for decisionmaking via "knowledge discovery." Much exploratory data analysis and inferential statistics concern the same type of problems, so what is different about data mining? What is similar? In the course, I will attempt to answer these questions by providing a broad survey of the problems that motivate data mining and the approaches used to solve them.

August 3, 2008

CE_01C 8:30 a.m.–5:00 p.m.

Generalized Linear Mixed Models: Theory and Applications

ASA

Instructor(s): Oliver Schabenberger, SAS Institute Inc.; Walter Stroup, University of Nebraska-Lincoln

CE_07C 8:00 a.m.–12:00 p.m.

Design and Analysis of Epidemiologic Studies of Gene-Environment Interactions

Section on Statistics in Epidemiology

Instructor(s): Raymond Carroll, Texas A&M University; Nilanjan Chatterjee, National Cancer Institute

Most common human diseases have a multifactorial etiology involving a complex interplay of genetic and environmental exposures. Understanding how genetic and environmental exposures interact and jointly influence the risk of a complex disease can be important for both biological and public health purposes. We will present the state of the art of efficient design and analysis for studies of gene-environment interaction by statisticians, epidemiologists, and geneticists. Topics covered will include population- and family-based case-control designs, stratified sampling designs, modern semiparametric methods for analysis of case-control data, estimation of haplotype-environment interactions, and flexible modeling approaches to empirical Bayes methods. We will blend theory and applications with illustrations using real examples and software implementation.

CE_08C 8:30 a.m.–5:00 p.m.

Modern Practical Bayesian Clinical Trial Design

Section on Bayesian Statistical Science

Instructor(s): Peter F. Thall, University of Texas, M.D. Anderson Cancer Center; J. Kyle Wathen, University of Texas, M.D. Anderson Cancer Center

We will cover practical Bayesian methods for clinical trial design and conduct. Attendees should have a master's degree in statistics, or equivalent experience, and an understanding of elementary Bayesian concepts. There will be numerous illustrations using actual clinical trials. Drawn from oncology, examples will include methods for eliciting and calibrating priors, incorporating historical data, and using computer simulation to establish a design's frequentist properties. The morning will cover phase I designs-including dose-finding using the continual reassessment method and logistic regression, finding optimal dose pairs of a two-agent combination, and accommodating multiple toxicities-and phase II designs, including a paradigm for monitoring multiple discrete outcomes, randomized phase II trials, monitoring event times, hierarchical Bayesian methods for trials with multiple disease subtypes, and using regression to account for patient heterogeneity. The afternoon will cover phase I/II dose-finding based on efficacy-toxicity trade-offs, optimizing schedule of administration, jointly optimizing dose and schedule, adaptive randomization, a geometric approach to treatment comparison based on two-dimensional parameters, and designs to evaluate multistage dynamic treatment regimes.

CE_09C 8:30 a.m.–5:00 p.m.

Statistical Challenges in Proteomics

Biometrics Section

Instructor(s): Scott C. Schmidler, Duke University

Proteomics is the next frontier in the rapidly evolving field of bioinformatics. I will provide an introduction to the principal aims, technologies, and statistical issues arising in structural and functional proteomics studies. Topics will include experimental data sources (e.g., X-ray, NMR, mass spectrometry [MALDI, SELDI, MS/MS], peptide arrays), statistical problems in structural proteomics (e.g., molecular comparison and database search, classification of structures, structure-based function prediction), and statistical problems in functional proteomics (e.g., fragment identification, normalization and registration of spectra, peak finding, sample comparison, classification and biomarker identification).

CE_10C 8:30 a.m.–5:00 p.m.

Systematically Improving Your Professional Practice

Committee on Career Development

Instructor(s): Doug Zahn, Zahn and Associates

Participants will enhance their career/professional development by learning a process for running effective one-on-one consultations/collaborations/meetings. While this process is valuable, it is not enough to markedly improve your practice by itself. Doing this requires facing the uncomfortable fact that no matter how carefully we plan a meeting, breakdowns will occur as we interact with our colleagues, clients, students, supervisors, and staff. Thus, we will learn a process to recognize and resolve breakdowns effectively and efficiently. Attendees should have at least one year of professional practice.

CE_11C 8:30 a.m.–5:00 p.m.

Principles of Statistical Design

ASA

Instructor(s): George Casella, University of Florida

We will cover the principles and practice of statistical design, paying attention to the setup and implementation of an experiment and the underlying theory that allows valid inferences. The course will begin with a review of the basic tools for statistical design and the statistical package R. The more common designs (e.g., factorial completely randomized designs, randomized complete blocks) and their variations (e.g., Latin squares) will be covered. Emphasis will be on designing the experiment to obtain the best inference on treatment contrasts, and designs will be illustrated will real data problems. We will focus on microarray designs and spend a lot of time on split plots and their variations (e.g., strip plot, repeated measures). Finally, we will move to confounding (e.g., incomplete blocks, fractions). This course is aimed at professional-level statisticians or interested faculty and graduate students. Attendees should have a working knowledge of statistical methodology and data analysis (e.g., Rawlings et al. *Applied Regression Analysis*, Springer-Verlag, 1998).

CE_12C 1:00 p.m.–5:00 p.m.

Sampling in Networks

Section on Survey Research Methods

Instructor(s): Steven K. Thompson, Simon Fraser University

Network models are in increasing use to describe populations, including socially networked human populations, computer and communication networks, and gene regulatory networks. A network has nodes (e.g., people) and links (e.g., relationships between people). The nodes may have characteristics of interest, and the relationships may be of different types and strengths. Network data, however, generally represent a sample from the wider population network of interest. This short course will cover methods for obtaining samples from networks and using the sample data to make inference about characteristics of the population network. In many cases the only practical way to obtain a large enough sample from the population is to follow links from sample individuals to add more individuals to the sample. For example, in studies of the risk behaviors in people at risk for HIV/AIDS, the population is hidden so standard sampling designs cannot be applied. Instead, researchers follow social referrals from individuals in the sample to find more members of the hidden population. Similarly, in studies of the World Wide Web, links or connections from sites in the sample are followed to add more sites to the sample. Network methods also turn out to be useful for spatial sampling in environmental and ecological sciences where the populations tend to be highly clustered or rare. Link-tracing sampling designs will be described, together with design-based and Bayes methods for estimating population characteristics based on such samples. Computational methods and available software also will be described.

August 4, 2008

CE_13C 8:00 a.m.–12:00 p.m.

Evaluating Probability of Success for Internal Decisionmaking in Early Drug Development

Biopharmaceutical Section

Instructor(s): Narinder Nangia, Abbott Laboratories; Martin King, Abbott Laboratories; Jane Qian, Abbott Laboratories

Early development (the "learning stage") is a crucial period of the drug development process, as decisions to continue or halt development of a compound must be made with incomplete information. Relying solely on p-values from phase I-II studies for making drug development milestone decisions is an inefficient approach, as it ignores several important determinants of future success. We will discuss the statistical tools that enable quantification of the uncertainty associated with results coming from learning stage studies. These tools use the Bayesian approach to exploit the totality of accumulated data/knowledge in a formal way for internal decisionmaking in early drug development. Posterior and/or predictive probabilities computed in a Bayesian paradigm are easy to interpret and provide much more relevant information than p-values for decisionmaking. We will also discuss evaluation of probability of a successful phase III trial through clinical trial simulations. Examples from the CNS, inflammation and oncology therapeutic areas will be considered for evaluation of probability of success for drug candidates in meeting the target product profile.

CE_14C 8:00 a.m.–12:00 p.m.

U-Statistics for Scoring Multivariate Data: From Sports to Genetics

ASA

Instructor(s): Knut M. Wittkowski, The Rockefeller University; Tingting Song, The Rockefeller University Center for Clinical and Translational Science

We will extend commonly used u-statistics for univariate and censored data to multivariate data with innovative applications in sports, economics, sociology, biology, and medicine. The course consists of four parts: stratification as a means to improve McNemar-type tests for trio data in genetics ('TDT') and adapt them to various genetic models; history of u-statistics; how information about relationships between variables can be incorporated through transforming data, converting data into partial orderings, and combining partial orderings; and computational and statistical aspects of screening studies involving thousands of variables (SNP or gene-expression microarrays) and nonparametric "factor analyses." Demonstrations will be based on spreadsheets, functions from muStat (available from http://cran.r-project.org and http://csan.insightful.com), and web services available from http://muStat.rockefeller.edu. Prerequisites: Basic knowledge of statistics and programming. Recommended Textbook: Lehmann, E.L. (1975). *Nonparametrics: Statistical Methods Based on Ranks*. Holden-Day.

CE_15C 8:30 a.m.–5:00 p.m.

Analysis of Clinical Trials: Theory and Applications

Biopharmaceutical Section

Instructor(s): Christy Chuang-Stein, Pfizer Inc; Alex Dmitrienko, Eli Lilly and Company; Keaven Anderson, Merck & Co., Inc.

We will cover analysis of stratified data, multiple comparisons and multiple endpoints, and interim analysis and interim data monitoring by presenting practical advice from experts, offering a well-balanced mix of theory and applications, and discussing regulatory considerations. The discussed statistical methods will be implemented using SAS software, and clinical trial examples will be used for illustration. This course is for statisticians working in the pharmaceutical or biotechnology industries, as well as contract research organizations. It is equally beneficial to statisticians working in institutions that deliver health care and government branches that conduct health care-related research. Attendees must have basic knowledge of clinical trials. Familiarity with drug development is highly desirable, but not necessary.

CE_16C 8:30 a.m.–5:00 p.m.

Graphics of Large Data Sets

Section on Statistical Graphics

Instructor(s): Antony Unwin, Augsburg University; Heike Hofmann, Iowa State University

Graphics are great for exploring data, but how can they be used for looking at the large data sets commonplace today? Large data sets bring new complications and require different emphases and approaches. In this course, based on Graphics of Large Datasets, we will discuss how to look at ways of visualizing large datasets, whether large in number of cases, number of variables, or both. Data visualization is useful for data cleaning, exploring data, identifying trends and clusters, spotting local patterns, evaluating modeling output, and presenting results. It is essential for exploratory data analysis and data mining. Data analysts, statisticians, and computer scientists should benefit from attending this course. Participants are welcome to bring laptops and should have knowledge of standard statistical graphics and experience carrying out data analysis. Either the software Mondrian (which can be downloaded from *stats.math.uni-augsburg.de/Mondrian/*) or, if you use R, the R package iPlots should be installed.

CE_17C 8:30 a.m.–5:00 p.m.

Statistical Evaluation of Medical Tests and Biomarkers for Classification

Section on Statistics in Epidemiology

Instructor(s): Margaret S. Pepe, Fred Hutchinson Cancer Research Center; Holly Janes, Fred Hutchison Cancer Research Center; Todd A. Alonzo, University of Southern California Keck School of Medicine

Development of biomarkers and medical diagnostic devices has accelerated. Their rigorous evaluation is a high priority for research, yet principles and techniques for the design and analysis of these studies are not widely known. There are fundamental differences among methods for therapeutic and etiologic studies. Moreover, much basic methodology has developed recently. We will cover estimation and comparison of Receiver Operating Characteristic (ROC) curves and describe extensions to adjust for covariates that affect biomarker/test measurements. For assessing factors associated with test performance, ROC regression methods will be presented. We also will consider how to evaluate the benefit of a new test when standard tests or clinical variables exist. Second, we will consider the design of case-control studies most common in this field. Sample size calculations and optimal choice of case-control ratio will be presented and the attributes and limitations of matching controls to cases will be discussed. Third, prospective studies will be considered. Finally, we will discuss problems incurred when the gold standard reference test is, itself, subject to error. A suite of freely available Stata programs will implement analyses. Prerequisite: introductory statistics.

CE_18C 8:30 a.m.–5:00 p.m.

Computational Statistics: Methods for Monte Carlo Integration and Optimization

Section on Statistical Computing

Instructor(s): Jennifer A. Hoeting, Colorado State University; Geof H. Givens, Colorado State University

This course will consist of two parts: a morning session on Monte Carlo integration strategies and an afternoon session on optimization methods. We will survey a variety of techniques, ranging from classic to state-of-the-art. The course will be based on Computational Statistics, and is aimed at quantitative scientists and statisticians who are unfamiliar with these methods. Upper division undergraduate mathematical literacy is recommended. Many problems in statistics require the evaluation of integrals that cannot be solved analytically, particularly in Bayesian statistics. We will cover Monte Carlo integration, importance sampling and variance reduction techniques, and Markov chain Monte Carlo methods. Optimization also plays a central role in statistics, particularly in numerical maximum likelihood estimation. The afternoon session will cover Newton-like methods, Gauss-Seidel iteration, tabu algorithms, simulated annealing, genetic algorithms, and the EM algorithm and its variants. We seek to give students a practical understanding of how and why existing methods work, enabling them to use modern statistical methods effectively. We focus on methodological concepts, and not details of computer programming. Examples are drawn from diverse fields including bioinformatics, ecology, and medicine.

CE_19C 1:00 p.m.–5:00 p.m.

Methods for Identifying and Dealing with Overdispersed Regression Models

ASA

Instructor(s): Joseph M. Hilbe, Arizona State University

We will define overdispersion in the context of binomial and count models and specify the difference between apparent and real overdispersion and how to identify each. We also will show methods that can be used to eradicate apparent overdispersion from a model, as well as discuss methods used to deal with real overdispersion.

CE_20C 1:00 p.m.–5:00 p.m.

Adaptive Analysis of Data: Tests of Significance and Confidence Intervals

ASA

Instructor(s): Tom O'Gorman, Northern Illinois University

I will present several adaptive methods for the analysis of data, beginning with a two-sample adaptive test, and then present an adaptive method of testing any subset of coefficients in a multiple regression model. I will also describe adaptive tests for interaction and main effects in the analysis of factorial experiments and adaptive tests for slope. The advantage of adaptive tests is that they are usually more powerful than the traditional tests for non-normal error distributions. As there is little power loss with normal error distributions, adaptive tests can be recommended for general use in studies having more than 20 observations. For each adaptive test, we will compare its performance to the traditional method, and I will show how to perform the test using a SAS macro. Adaptive tests used in the analysis of repeated measurements will be described and compared to the nonadaptive mixed model tests. In addition, I will describe a method of computing adaptive confidence intervals. Attendees should be familiar with basic statistical modeling-including multiple regression and the analysis of variance-at the level of *Applied Regression Analysis* (1998) by Draper and Smith.

August 5, 2008

CE_21C 8:00 a.m.–12:00 p.m.

Analysis of Multivariate Failure Time Data

ASA

Instructor(s): Danyu Lin, The University of North Carolina at Chapel Hill

Multivariate failure time data arise when each study subject can potentially experience multiple events or when there exists clustering of subjects such that failure times within the same cluster are correlated. Major complications in analyzing such data include the dependence among related failure times and censoring due to limited follow-up or competing events. This short course presents a variety of statistical models and methods for the analysis of multivariate failure time data. We will discuss marginal and frailty models, paying primary attention to semiparametric regression methods. Relevant software will be described, and a number of clinical and epidemiologic studies will be provided for illustrations.

CE_22C 8:00 a.m.–12:00 p.m.

Fundamental Statistics Concepts in Presenting Data: Principles for Constructing Better Graphics

Section on Statistical Graphics

Instructor(s): Rafe Donahue, Vanderbilt University

Data displays are mental models for understanding distributions of data. At the heart of any data display lays the distribution of the data; a model for this distribution includes demonstrating and exposing sources of variation in the distribution. Like a good map, a display of data ought to operate on several levels. At the lowest level (the highest level of granularity) are the data, themselves. Further up are the actual distributions, each with its component summaries, such as the mean or relevant quantiles. At the highest level are sources of variation in these distributions, the parameters in the (mental) model for understanding the data. The closer an architect can come to showing all these levels, the more information will be conveyed. I will present a number of principles, both developed by the masters (e.g., Minard, Tufte, Cleveland, Wilkinson, Wainer) and discovered by me, for constructing displays that will allow the architect of the data display to present the data for improved understanding; it will not be a "Don't use pie charts" or "Here's a bad graph from *USA Today*" course. We will focus on uncovering and formulating principles for presenting data visually. Examples will abound.

CE_23C 8:30 a.m.–5:00 p.m.

Bayesian Methods and Software for Data Analysis

Section on Bayesian Statistical Science

Instructor(s): Bradley P. Carlin, University of Minnesota; Thomas A. Louis, Johns Hopkins University

This course will introduce hierarchical and empirical Bayes methods, demonstrate their usefulness in challenging applied settings, and show how they can be implemented using modern Markov chain Monte Carlo (MCMC) computational methods. We will provide an introduction to and live demonstration of WinBUGS, the most general Bayesian software package available to date, and BRugs, a convenient function for calling BUGS from R. Use of the methods will be demonstrated in advanced high-dimensional model settings (e.g., nonlinear longitudinal modeling or spatiotemporal estimation and mapping), where the MCMC Bayesian approach often provides the only feasible alternative incorporating all relevant model features. Participants should have an MS (or advanced undergraduate) understanding of mathematical statistics at the Hogg and Craig (1978) or Casella and Berger (2001) level. Basic familiarity with common statistical models (e.g., the linear regression model) and computing will be assumed, but we will not assume significant previous exposure to Bayesian methods or Bayesian computing. This course is aimed at students and practicing statisticians who are intrigued by all the fuss about Bayes and Gibbs, but who may still mistrust the approach as theoretically mysterious and practically cumbersome.

CE_24C 8:30 a.m.–5:00 p.m.

Models for Discrete Repeated Measures

ASA

Instructor(s): Geert Verbeke, K.U. Leuven; Geert Molenberghs, Hasselt University

Starting from a brief introduction to the linear mixed model for continuous longitudinal data, we will formulate extensions to model outcomes of a categorical nature, including counts and binary data. Based on Verbeke and Molenberghs (2005), several families of models will be discussed and compared, from an interpretational and computational point of view. First, we will discuss models for the full marginal distribution of the outcome vector. Such models allow inference to be based on maximum likelihood principles, but they have the disadvantage of requiring complete specification of all higher-order interactions. We will talk about two alternatives: random-effects models and semiparametric marginal models with specification of the first moments only, or the first and second moments only. We will discuss and illustrate in full detail estimation and inference, and we will extensively argue that both approaches yield parameters with completely different interpretations. Finally, when analyzing longitudinal data, one is often confronted with missing observations. We will show that, if no appropriate measures are taken, missing data can cause seriously biased results and interpretational difficulties. Methods to properly analyze incomplete data, under flexible assumptions, will be presented and key concepts of sensitivity analysis will be introduced.

CE_25C 8:30 a.m.–5:00 p.m.

Mixed Models for the Practicing Statistician

Section on Statistics and the Environment

Instructor(s): Linda J. Young, University of Florida; Ramon C. Littell, University of Florida

Data sets from designed experiments, sample surveys, and observational studies often contain correlated observations due to random effects and repeated measures. Mixed models can be used to accommodate the correlation structure, produce efficient estimates of means and differences between means, and provide valid estimates of standard errors. Repeated measures and longitudinal data require special attention because they involve correlated data that arise when the primary sampling units are measured repeatedly over time or under different conditions. We will use normal theory models for random effects and repeated measures ANOVA to introduce the concept of correlated data. We will then extend these models to generalized linear mixed models for the analysis of non-normal data, including binomial responses, Poisson counts, and over-dispersed count data. We will discuss methods of assessing the fit and deciding among competing models. Radial smoothing splines can be represented as mixed models, and we will illustrate their application. We will illustrate PROC GLIMMIX in the SAS system using practical examples from pharmaceutical trials, environmental studies, educational research, and laboratory experiments.

CE_26C 8:30 a.m.–5:00 p.m.

Multiple Imputation of Missing Data

ASA

Instructor(s): Paul D. Allison, University of Pennsylvania

This course will cover both the conceptual foundations and practical details of implementing multiple imputation. Conventional methods for handling missing data typically yield biased estimates and/or incorrect standard errors. By contrast, multiple imputation produces estimates with nearly optimal properties under weaker assumptions. I will explain the assumptions of "missing at random" and "missing completely at random." After a brief review of conventional methods, we will consider multiple imputation based on linear regression with random draws. We will examine implementation using the MCMC algorithm in SAS PROC MI in detail, and then move on to the role of the dependent variable, imputation under a restricted range, imputation of categorical variables, multivariate inference, interactions and nonlinearities, congeniality of data model and imputation model, longitudinal data, nonignorable missing data, and imputation by chained equations (demonstrated using the ice command in Stata).

CE_27C 1:00 p.m.–5:00 p.m.

Meta-Analysis: Statistical Methods for Combining the Results of Independent Studies

ASA

Instructor(s): Ingram Olkin, Stanford University

Meta-analysis enables researchers to synthesize the results of a number of independent studies designed to determine the effect of an experimental protocol, such as an intervention, so the combined weight of evidence can be considered and applied. Increasingly, meta-analysis is being used in the health sciences, education, and economics to augment traditional methods of narrative research by systematically aggregating and quantifying research literature. The information explosion in almost every field coupled with the movement toward evidence-based decisionmaking and cost-effective analysis has served as a catalyst for the development of procedures to synthesize the results of independent studies. In this course, I will provide a historical perspective of meta-analysis and discuss some of its issues. The statistical methodology will include discussions of nonparametric and parametric models, effect sizes for proportions, fixed versus random effects, regression, and ANOVA models. New material on multivariate models also will be presented.

CE_28C 1:00 p.m.–5:00 p.m.

Analysis of Censored Health Outcomes Data: Developments for the Last 10 Years

Biopharmaceutical Section, Section on Teaching Statistics in the Health Sciences

Instructor(s): Heejung Bang, Weill Medical College of Cornell University; Hongwei Zhao, University of Rochester

Medical cost and quality-adjusted lifetime are common health outcomes data from clinical trials and observational studies. Although these data look different, they share many statistical properties and can be understood in a unified framework. Just like standard survival data, censoring is an important issue in these data. Despite the analogy, censoring mechanism is informative, different from the traditional paradigm. It has been a decade since it was shown that the use of most standard statistical techniques (e.g., sample mean, linear regression, and Kaplan-Meier estimator) can be invalid. However, we often find that even experienced researchers still use traditional methods for the analysis of health outcome data in practice. In this course, we will review valid methods for statistical estimation and inference that have been developed in last 10 years. Unfortunately, not all are easy or user-friendly, and no commercial software is available so far. Therefore, we will suggest methods as practical solutions for practitioners. We also will present the analytic relationships among well-known medical cost estimators recently identified. Extended applications to customer lifetime value and cost-effectiveness analysis will be discussed. Course prerequisite is basic knowledge of survival analysis.

August 6, 2008

CE_29T 8:00 a.m.–9:45 a.m.

Meta-Analysis: Concepts and Applications

ASA

Instructor(s): Michael Borenstein, Biostat, Inc.; Hannah R. Rothstein, Biostat, Inc.

Meta-analysis is a set of statistical procedures to synthesize data from multiple studies. When the studies share a common effect size, the meta-analysis yields a more precise estimate of that effect than any single study, and when the effect varies from one study to the next, meta-analysis may be used to explain the variation. Meta-analyses are used to inform policy, obtain approval for drugs, and design research. They also play a key role in grant applications and publications. We will explain the concept of meta-analysis and show how to compute treatment effects and a combined effect, assess heterogeneity, and explain variation in treatment effects across studies. We will discuss the difference between fixed and random effects models and address common criticisms of meta-analysis. We will demonstrate Comprehensive Meta Analysis Version 2, a program developed with funding from the NIH. This course is intended for people who perform or interpret meta-analyses. Attendees should have some familiarity with meta-analysis, but the course will cover the basics before moving on to advanced topics.

CE_30T 8:00 a.m.–9:45 a.m.

Determining Sample Size and Power in Study Planning: nQuery Advisor 7.0

ASA

Instructor(s): Janet Elashoff, Statistical Solutions; Brian Sullivan, Statistical Solutions

Choosing an adequate sample size is a vital part of study planning. We will review statistical methods for determining study sample size and power. Using nQuery Advisor with real examples, we will demonstrate the steps in sample size determination from specifying the design to writing a sample size justification statement. We will provide tips for the toughest problem in sample size determination-eliciting the information needed to specify "effect" sizes and "guesstimate" standard deviations-and we encourage discussion. We will illustrate sample size planning for survival studies with user specified hazard ratios and illustrate the effects of accrual and dropout patterns on required sample size. We will show the relationships between sample size methods for tests, confidence intervals, and noninferiority and equivalence studies. We will discuss the logistic and power issues of unequal n's and stratification and show how to make the last step in study planning, the creation of randomization lists, easy. Attendees should be experienced in the use of data analysis methods commonly taught in master's programs in statistics.

CE_31T 8:00 a.m.–9:45 a.m.

An Introduction to Stat Studio for SAS/STAT Users

ASA

Instructor(s): Rick Wicklin, SAS Institute Inc.

Stat Studio 3.1 is new statistical software in SAS 9.2. It provides a highly flexible programming environment in which you can run SAS/STAT and SAS/IML analyses and display the results with dynamically linked graphics and data tables. You can also call SAS procedures from an IML program. Stat Studio is intended for data analysts who write SAS programs to solve statistical problems but need more versatility for data exploration and model building. This workshop introduces Stat Studio to SAS/STAT users. You will learn how to use the point-and-click features of Stat Studio for analyzing data interactively, write programs that use interactive graphics to display diagnostic statistics computed by SAS/STAT procedures for model assessment and outlier identification, and write programs that implement modern statistical methods, such as bootstrap algorithms and nonparametric smoothing techniques. Attendees should have basic knowledge of SAS/STAT procedures such as FREQ, REG, and LOGISTIC. Experience with SAS/IML and object-oriented programming is helpful, but not required.

CE_32T 8:00 a.m.–9:45 a.m.

From Software to Solutions in Statistics and Risk Analysis

ASA

Instructor(s): Shawn Harahush, Palisade

The world of business and education has become more complex with the decision of the type of software a business will use to successfully manage their incoming data. Palisade, a world leader in risk analysis, has been creating software solutions for more than 20 years. Palisade's flagship product, @RISK, integrates into Microsoft Excel to provide a powerful Monte Carlo simulation engine to the ease-of-use environment. Palisade's StatTools also integrates with Excel to provide reliable and easy-to-use statistics to an easily accessible program. NeuralTools adds sophisticated neural networks analysis into an easy-to-use and familiar interface: Microsoft Excel.

CE_33T 10:00 a.m.–11:45 a.m.

EastAdapt: A Module for Late-Stage Adaptive Trial Design Within the EastÆ 5 Software System

ASA

Instructor(s): Cyrus R. Mehta, Cytel Inc.

We will demonstrate EastAdaptÆ, a major upgrade of the adaptive design module of East that is used for designing and simulating late-stage (phase II and phase III) clinical trials. EastAdapt makes it possible to design clinical trials with a data-dependent mid-course correction to sample size, spending function, and number of future interim analyses and their spacing without inflating the type I error. EastAdapt's simulations are used to determine the operating characteristics of the adaptive design and compare them to those of a classical group sequential design. A major new capability is the ability to compute valid p-values, point estimates, and confidence intervals at the end of the adaptive clinical trial. Another major new EastAdapt capability is the ACR Method for performing the adaptive hypothesis test. With the ACR method, one can use the usual sufficient statistic, rather than the Cui, Hung and Wang (1999) weighted statistic to determine statistical significance.

CE_34T 10:00 a.m.–11:45 a.m.

Survey Data Analysis with Stata

ASA

Instructor(s): Jeffrey Pitblado, Stata Corp LP

This workshop will cover how to use Stata for survey data analysis assuming a fixed population. Knowledge of Stata is not required, but attendees should have some statistical knowledge, such as what is typically covered in an introductory statistics course. We will begin by reviewing the sampling methods used to collect survey data and how they affect the estimation of totals, ratios, and regression coefficients. We will then cover the three variance estimators implemented in Stata's survey estimation commands. Stata with a single sampling unit, certainty sampling units, subpopulation estimation, and poststratification will also be covered. Each topic will be illustrated with an example in a Stata session.

CE_35T 10:00 a.m.–11:45 a.m.

Nonparametric Regression Modeling in SAS Software

ASA

Instructor(s): Weijie Cai, SAS Institute Inc.

Nonparametric modeling is widely employed in modern statistical analysis in cases where only limited knowledge of the underlying model is available. You can use nonparametric modeling to discover nonlinear dependencies in your data, enabling you to develop parsimonious parametric models. This workshop is intended for a broad audience of statisticians and data analysts who are interested in nonparametric regression modeling. In it, I will describe methods and SAS tools for fitting local regression models with the LOESS procedure, penalized spline models with the TRANSREG procedure, thin-plate spline models with the TPSPLINE procedure, generalized additive models with the GAM procedure, penalized spline and radial basis function models using a mixed model approach with the GLIMMIX procedure, and selected basis functions models with the GLMSELECT procedure. The audience should have a basic understanding of regression theory.

CE_36T 10:00 a.m.–11:45 a.m.

Introduction to CART: Data Mining with Decision Trees

ASA

Instructor(s): Mikhail Golovnya, Salford Systems

This course, intended for the applied statistician wanting to understand and apply the CART methodology for tree-structured nonparametric data analysis, will emphasize practical data analysis involving classification. All concepts will be illustrated using real-world examples. The course will begin with an intuitive introduction to tree-structured analysis. Working through examples, we will review how to read CART output and set up basic analysis. This session will include performance evaluation of CART trees and cover ways to search for possible improvements of the results. Once a basic working knowledge of CART has been mastered, we will focus on critical details essential for advanced CART applications, including choice of splitting criteria, choosing the best split, using prior probabilities to shape results, refining results with differential misclassification costs, the meaning of cross validation, tree growing, and tree pruning. The course will conclude with discussion of the comparative performance of CART versus other computer-intensive methods, such as artificial neural networks and statistician-generated parametric models.

CE_37T 1:00 p.m.–2:45 p.m.

New Software for the Design, Analysis, and Reporting of Bioequivalence and Clinical Pharmacology Trials

ASA

Instructor(s): Yannis Jemiai, Cytel Inc.; Pralay Senchuadhuri, Cytel Inc.

Cytel Inc. introduces a software package for the design, analysis, and reporting of early phase clinical pharmacology trials, as well as pivotal and nonpivotal bioequivalence trials. Key development members from Cytel Inc. will demonstrate how to quickly design parallel and crossover clinical trials with superiority, noninferiority, or equivalence objectives; create, import, and explore data sets; produce and compare analyses and plots; and construct standardized templates to generate standardized reports of your work

CE_38T 1:00 p.m.–2:45 p.m.

New Procedures and Features for Clustered and Survey Data Analysis in SUDAAN Release 10

ASA

Instructor(s): Angela Pitts, RTI International; George G. Brown, RTI International

This workshop will highlight two new procedures and several new features in SUDAAN Release 10, which will be available in early August 2008. SUDAAN is a statistical software package for the analysis of complex survey and other cluster-correlated data. We will focus on the new PROC HOTDECK procedure that conducts sequential weighted hot deck imputation; the PROC WTADJUST procedure that computes weight adjustments; and the addition of model-adjusted risk ratios, a test for the proportional odds assumption, exponentiated point estimates defined by EFFECTS statement contrasts, a SORTED option on the NEST statement, the use of character variables in all procedures, and several enhancements to the PRINT statement. The workshop will include a brief introduction to SUDAAN syntax. Attendees are not required to be SUDAAN users, but should have knowledge of statistical issues that arise when analyzing survey and other correlated data. The new SUDAAN features will be demonstrated on complex survey data. We will demonstrate proper implementation of SUDAAN, provide interpretation of the output, and discuss statistical issues related to the data. All course material, including a 30-day trial version of SUDAAN Release 10, will be provided.

CE_39T 1:00 p.m.–2:45 p.m.

Introduction to Bayesian Analysis Using SAS Software

ASA

Instructor(s): Fang Chen, SAS Institute Inc.

Bayesian methods have become increasingly popular in recent years in a number of disciplines. This workshop will provide an introduction to Bayesian methods with applications in the generalized linear model and survival analysis. The first part will provide an overview of Bayesian methodology, including motivation and Bayesian inference, and computational methods and convergence diagnostics relevant to the SAS implementation. The second part will cover applications using new capabilities in SAS/STAT software in the GENMOD, LIFEREG, and PHREG procedures, which are based on Gibbs sampling. Examples will include linear regression, logistic regression, Poisson regression, Cox regression, parametric survival models, and the piecewise exponential model. Note that these enhanced procedures are available in the newly available SAS 9.2. A master's-level knowledge of statistics is assumed, as well as experience with generalized linear models and survival analysis. Previous exposure to Bayesian methods is useful, but not required.

CE_40T 1:00 p.m.–2:45 p.m.

Introduction to MARS: Predictive Modeling with Nonlinear Automated Regression Tools

ASA

Instructor(s): Mikhail Golovnya, Salford Systems

This workshop will introduce the main concepts behind Jerome Friedman's MARS, a modern regression tool that can help analysts quickly develop superior predictive models. MARS is a nonlinear automated regression tool that can trace complex patterns in data. It automates the model specification search, including variable selection, variable transformation, interaction detection, missing value handling, and model validation. Conventional regression models typically fit straight lines to data. Although this usually oversimplifies the data structure, the approximation is sometimes good enough for practical purposes. However, in the frequent situations in which a straight line is inappropriate, an expert modeler must search tediously for transformations to find the right curve. MARS approaches model construction more flexibly, allowing for bends, thresholds, and other departures from straight lines from the beginning. Attendees will be presented with MARS' key benefits.

CE_41T 3:00 p.m.–4:45 p.m.

Exact Methods Module for EastÆ 5: Design, Simulate, Analyze, and Monitor Binomial Endpoint Trials by Exact Inference Methods

ASA

Instructor(s): Anthiyur Kannappan, Cytel Inc.; Pralay Senchuadhuri, Cytel Inc.

We will present a newly added special module, "Exact Methods" for East 5, to design, simulate, analyze, and monitor binomial endpoint trials by exact inference methods. This module includes procedures for Simon's two-stage optimal, one sample (group sequential), paired proportions, two sample superiority (difference, ratio, Fisher's), two sample noninferiority (difference, ratio), and two sample equivalence. This module will be especially suitable to situations where the sample sizes are not expected to be large. The usual features of EastÆ-boundary chart, enhanced simulation, and interim monitoring capability-are available for the procedures in this module. The additional feature is the ability to input 2x2 data in the interim monitoring sheet and get the exact inference method results there.

CE_42T 3:00 p.m.–4:45 p.m.

Structural Analysis of Time Series Using the SAS/ETS UCM Procedure

ASA

Instructor(s): Rajesh Selukar, SAS Institute Inc.

This workshop will introduce the SAS/ETS UCM procedure, which enables analysis of time series data by using structural models. Structural models provide regression-like decomposition of the response series into components such as trend, seasonal or other periodic, and linear and nonlinear regression effects. Apart from the series forecasts, this methodology provides estimates of these unobserved components, which are useful in practical decisionmaking. Participants will learn to identify, diagnose, and use structural time series models for time series data in a variety of situations. The course will cover novel time series techniques, including approximation of long and complex seasonal patterns by using splines and incorporation of linear and nonlinear regression effects with time varying coefficients. Several real-life examples will be used to demonstrate the functionality of the UCM procedure. Participants also will learn the relationship between the ARIMA models-another class of models widely used for analyzing time series data-and structural models.

CE_43T 3:00 p.m.–4:45 p.m.

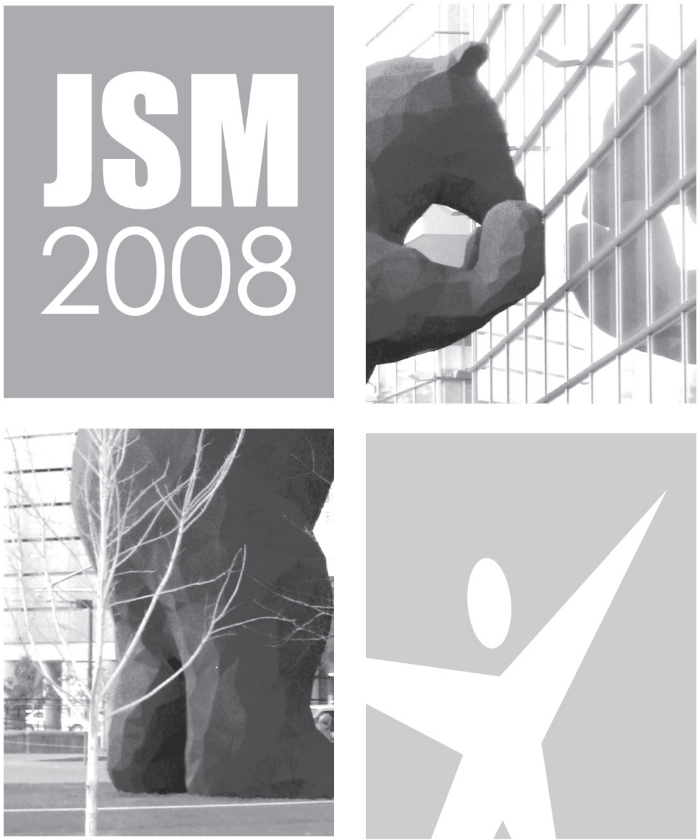
Advances in Data Mining: Jerome Friedman's TreeNet/MART and Leo Breiman's Random Forests

ASA

Instructor(s): Mikhail Golovnya, Salford Systems

This workshop will present Leo Breiman's Random Forests and Jerome Friedman's TreeNet/MART. Random Forests and MART/TreeNet are advances to classification and regression tree software, which enable the modeler to construct predictive models of extraordinary accuracy. Random Forest is a tree-based procedure that makes use of bootstrapping and random feature generation. In TreeNet, classification and regression models are built gradually through a potentially large collection of small trees, each of which improves on its predecessor through an error-correcting strategy. I will show how the software is used to solve real-world data mining problems, discuss theory and what is novel in the software, highlight implementation, compare the two methodologies, and show where the software fits in terms of other data mining software.

Index



INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF PARTICIPANTS

INDEX OF CE INSTRUCTORS